Copula-Based Precipitation Fields Estimation Combining Data From Radar, Gauge And Microwave Attenuation

Dissertation
zur Erlangung des Doktorgrades an der
Fakultät für Angewandte Informatik
der Universität Augsburg

vorgelegt von

Wei Qiu

aus
China

Master of Science (ESPACE), TU München

2012

Erstes Gutachten:	Prof. Dr. Harald Kunstman
Zweites Gutachten:	Prof. Dr. Jucundus Jacobeit
Weiteres Gutachten:	Prof. Dr. Karl-Friedrich Wetzel

Tag der mündlichen Prüfung: 30.07.2012

acknowledgements

In the first place, I want to express my sincere gratitude to Prof. Dr. Harald Kunstmann for offering me the opportunity to work on the challenging topic about using Copula-based approaches to estimated precipitation fields combining data from various observation devices. Without his enlightening instruction, impressive kindness and patience, I could not have completed my thesis. From numerous discussions with him, what I have learned is not only the scientific knowledge, but also the way of thinking about research, which is invaluable for my future academic career.

My special thanks go to Dr. Stefanie Vogl and Dr. Patrick Laux for their generous helps since I started to learn the Copulas. I also appreciate that they have been always available for discussions and answering questions during all these years.

I also want to thank Dr. Joerg Seltmann for providing me the radar data from MOHP, DWD. Otherwise, this research work can not be done successfully.

Thanks for the financial and technical supports from PROCEMA project and KIT.

Additionally, I would like to thank all the members of the groups of Prof. Dr. Harald Kunstmann in IMK-IFU, KIT, for the friendly and pleasant environment they have created. Among them are Benjamin Fersch, Christian Chwala, Christof Lorenz, Ganquan Mao, Jianhui Wei, Moussa Waongo, Michael Warscher and Sven Wagner.

Finally, I offer my regards and blessings to all of those who supported me in any respect during the completion of the thesis.

Abstract

Rain gauges are considered to provide the best available information about absolute point rainfall intensity at ground level but are limited in estimating the precipitation fields. Radar measured precipitation fields provide the spatial patterns aloft but are biased with respect to the absolute rain fall intensities. Recently, it was shown that the microwave link attenuation is a promising complement to the traditional devices such as gauge and radar. This dissertation contributes to the problem of how to estimate precipitation fields by assimilating information from gauge, radar and MW-link, combining their advantages.

Since the dependence structure between different precipitation observations is usually non-Gaussian, Copulas are applied to describe the dependence structure between observations from rain gauges and radar at the corresponding grid cells. As rain gauges are not available for each radar grid cell, two Copula-based approaches namely the Copula parameter map and interpolated Copula parameter field are used to model the spatial distribution of the dependence structure between gauge and radar positive pairs. Finally precipitation fields are simulated which retain the radar derived spatial patterns but are corrected for biases in their intensities.

From theoretical point of view, all Copula-based techniques require independent and identically distributed (*i.i.d.*) data as a pre-requisite which is often neglected. Therefore, in this dissertation, the sensitivity of the Copula-based approaches to the violation of the *i.i.d.* assumption is studied and the influence of the ARMA-GARCH transformation to the final estimated precipitation fields is investigated.

In addition to the pure dependence structure, the marginal distributions of the time series are another key aspect of each Copula model. Therefore, the temperature and altitude driven approach is developed to represent the spatial distribution of marginal distribution for rain gauges. Simulation results from various combinations of

the spatial dependence structures and marginal distributions are compared to reveal the advantages and disadvantages of the different approaches.

The microwave link attenuation, measuring the line integrated precipitation at near-ground level, can be either directly included in the Copula-based approach or used to adjust the radar derived rainfall fields first. It is proven that, by integrating the observations from MW-links, the estimated precipitation fields are further improved, leading to the better simulations of the precipitation fields at the near-ground level.

The performance of the Copula-based data assimilation approaches is demonstrated for the Bavarian Alps and Alpine Forelands. The simulated precipitation fields are compared to the interpolated gauge fields (Ordinary Kriging) and also cross-validated with the available 31 rain gauges at grid scale, as well as the operationally corrected radar precipitation (Radolan). The Copula-based approaches perform similarly well as indicated by different validation measures and successfully estimate precipitation fields by combining data from various observation sources.

Zusammenfassung

Regenmessungen von meteorologischen Stationen gelten als die besten verfügbaren Informationen über Absolutwerte des Niederschlages in Bodennähe, sind aber nur begrenzt in der Lage die räumliche Niederschlagsverteilung abzuschätzen. Die von Radarmessungen abgeleiteten Niederschlagsfelder geben diese räumliche Verteilung realistisch wieder, sind aber in Bezug auf die absoluten Intensitäten mit Fehlern behaftet. Kürzlich wurde gezeigt, dass Niederschlagsinformation, die aus der Dämpfung von Mikrowellen Links abgeleitet werden kann eine vielversprechende Ergänzung zu den traditionellen Geräten wie Regenradar und Regentopf ist. Die vorliegende Arbeit beschäftigt sich mit dem Problem, realistische Niederschlagsfelder durch Kombination der verfügbaren Datenquellen (Radar, Regentopf und Mikrowellenlink) so abzuschätzen, dass die Vorteile des jeweiligen Meßverfahrens erhalten bleiben.

Da die Abhängigkeitsstruktur zwischen verschiedenen Niederschlagsbeobachtungen in der Regel nicht Gauß-verteilt ist, werden Copulas angewandt, um die Abhängigkeitsstruktur zwischen Regentopf und Radarmessung in einer Radargitterzelle zu beschreiben. Da jedoch nicht für jede Radargridzelle im Untersuchungsgebiet auch Regentöpfe vorhanden sind, werden zwei Copula-basierte Verfahren entwickelt, die es ermöglichen auch diese Gridzellen zu berücksichten. Das erste Verfahren basiert auf der Nutzung der sogenannten Copula parameter maps (CPM). Im zweiten Verfahren werden interpolierte Copula Parameter-Felder (ICPF) verwendet, um die räumliche Verteilung der Abhängigkeitsstrukturen zwischen Regentopf- und Radarmessung (positive Paare) zu modellieren. Schließlich werden mit Hilfe der entwickelten Copula-basierten Methoden Niederschlagsfelder simuliert, in denen die vom Radar gemessene räumliche Verteilung des Niederschlages beibehalten wird während Fehler in den Absolutwerten erfolgreich korrigiert werden können.

Aus theoretischer Sicht benötigen alle Copula-basierten Verfahren sogenannte independent and identically distributed (*i.i.d.*) Daten als Voraussetzung, was oft vernachlässigt wird. Daher wird in dieser Dissertation die Sensitivität Copula-basierter Methoden in Bezug auf die Verletzung der *i.i.d.* Annahme und der Einfluss der ARMA-GARCH-Transformation auf die simulierten Niederschlagsfelder untersucht.

Neben der reinen Abhängigkeitsstruktur sind die Randverteilungen der Zeitreihen ein weiterer wichtiger Aspekt eines jeden Copula-Modelles. Daher wird ein Ansatz entwickelt, bei dem die Temperatur-und Höhenabhängigkeit der Parameter dieser Verteilungen genutzt wird um eine Verbesserung des Copula-Modelles in Bezug auf die räumliche Anwendung zu erreichen. Simulationsergebnisse aus verschiedenen Kombinationen der räumlichen Abhängigkeitsstrukturen und Randverteilungen werden schließlich verglichen um die Vor- und Nachteile der jeweiligen Methode abschätzen zu können.

Die Mikrowellen-Dämpfung, die eine linienintegrierte Messung in Bodennähe darstellt, kann entweder direkt in den Copula-Ansatz einbezogen werden oder verwendet werden, um die aus Radarmessungen abgeleiteten Niederschlagsfelder zuerst anzupassen. Es kann gezeigt werden, dass durch die Integration der Beobachtungen von MW-Links, die geschätzten Niederschlagsfelder weiter verbessert werden, insbesondere im Ver- gleich zu bodennahen Messungen.

Die Performanz der Copula-basierten Verfahren zur Datenassimilation wird am Beispiel eines Untersuchungsgebietes demonstriert, das sich in den Bayerischen Alpen und im Alpenvorland befindet. Die simulierten Niederschlagsfelder werden mit interpolierten Feldern (Ordinary Kriging, 31 Meßstationen des DWD) verglichen und mit den zur Verfügung stehenden 31 Meßstationen validiert. Desweiteren wird ein Vergleich mit operationell korrigierten Radarmessungen durchgeführt (Radolan-Verfahren, entwickelt vom DWD). Es kann schließlich gezeigt werden, dass die Copula-basierten Ansätze eine zum Radolan-Verfahren vergleichbare Performanz aufweisen und geeignet sind, Informationen aus verschiedenen Meßverfahren so zu verbinden, dass sowohl die räumliche Niederschlagsverteilung als auch die Niederschlagsintensität realistisch abgeschätzt werden können.

Contents

List of Figures

List of Tables

Glossary

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CDF	Cumulative Distribution Function
CPM	Copula Parameter Map
CSMD	Convective and Stratiform dependent Marginal Distribution
DEM	Digital Elevation Model
DWD	Deutscher Wetterdienst or German Weather Service
DSD	Drop Size Distribution
FMD	Fixed Marginal Distribution
GoF	Goodness of Fit
ICPF	Interpolated Copula Parameter Field
IDW	Inverse Distance Weighted
i.i.d.	Independent and Identically Distributed
MAE	Mean Absolute Error
MD	Marginal Distribution
MLE	Maximum Likelihood Estimation
MOHP	Meteorological Observatory Hohenpeissenberg
MW-link	Microwave Link
NaN	Not A Number
NSE	Nash-Sutcliffe Efficiency
UTC	Coordinated Universal Time
PCA	Principle Component Analysis
PDF	Probability Density Function
r	Pearson's Correlation Coefficient
RMSE	Root Mean Square Error
TAMD	Temperature and Altitude driven Marginal Distribution
w/	With
w/o	Without

A	Attenuation
A_r	The elective aperture (area) of the receiving antenna
R	Radar derived rainfall
Ra	Range
F	Pattern propagation factor
G	Gauge derived rainfall
G_t	The gain of the transmitting antenna
L	MW-link derived rainfall
MWS	Mean wind speed
o	Observation values
P	Precipitation
P_t	The transmitter power of radar
RH	Relative humidity
T	Temperature
Z	Radar reflectivity
z	Altitude
θ_F	Frank Copula parameter
θ_G	Gumbel Copula parameter
θ_C	Clayton Copula parameter
σ	The radar cross section, or scattering coefficient, of the target
λ	The parameter of Exponential distribution

1

Introduction

1.1 Motivation

Sustainable water management has become an issue of major concern over the past decades (e.g. Pahl-Wost et al., 2007). The importance of water resources and management has also be highlighted in recent climate studies with respect to political decision making and risk analyses (e.g. IPCC, 2007; IPCC, 2008). Increasing water demands in agriculture, industry and households lead to the declining fresh water availability which restricts economic development and wealth particular in water scarce environments with weak or vulnerable infrastructures (e.g. Orr et al., 2009). Understanding the hydrological cycle is fundamental to investigate all things related to water, which is also a key to the proper management of water resources. Precipitation is the principle source of the Earth's water supply so that the spatio-temporal distribution of rainfall is an important aspect within hydrological processes (e.g. Kuchment et al., 2004).

Traditional rain gauges are good point local observation tools but limited to their low spatial representativeness. In contrast, weather radar reflectivity measurements can provide spatial pattern information. However, the transformation of reflectivity to rain rate is accompanied by tremendous uncertainties (e.g. Habib et al., 2008; Villarini et al., 2008). The differences between radar and rain gauge observations can amount to 100 percent or even more due to random and systematic errors as e.g. reported by Smith et al. (1996). The discrepancies between rain gauges and the corresponding radar measurements are mainly due to the different spatio-temporal sampling properties. Rain gauges measure ground precipitation only at one point in space, while radar

measures indirect volumetric rainfall aloft in the atmosphere depending on the radar elevation angle and range from the radar station (e.g. Ehret et al., 2002).

These traditional rainfall measurement techniques inspired the use of microwave link attenuation, which has been proven to provide accurate line integral rainfall observations at the near-surface level (e.g. Messer et al., 2006; Leijnse et al., 2008). Comparing to gauge and radar, the installation of MW-links designating for rainfall monitoring is costly and not complicated. Furthermore, it is also feasible to use of MW-link attenuation data from commercial cellular network operators for monitoring rainfall fields (e.g. Giuli et al., 1991, 1999, and 2007; Zinevich et al., 2008).

So, the challenge lies on how to assimilate the precipitation information from different observation devices, such as rain gauge, radar and MW-link. As the distribution of precipitation is usually non-Gaussian, Copula-based approaches are applied to assimilate data from rain gauge and MW-link to radar precipitation fields simultaneously by the studying the dependence structures among them.

1.2 State of the Art

Traditionally, without the spatial precipitation information from remote sensing techniques, the rain gauge observations alone were used to derive the precipitation fields by using various interpolation methods. Among them, nearest neighbour (e.g. Isaaks et al., 1989), inverse distance weighting, regression models (e.g. Bourrough and McDonell, 1998), trend surface analyses (e.g. Colins and Bolstadt, 1996) and Splines (e.g. Hutchinson et al., 1998a and 1998b; Bourrough and McDonell, 1998) are some examples, as well as the Kriging approach with a large set of sub-methods developed (e.g. Isaaks et al., 1989; Bollerlslev et al., 1986; Goovaerts et al., 2000; Haberlandt et al., 2007). The mixture use of different methods can also been found such as regression combined with Kriging (e.g. Erxleben et al., 2002). All those methods are different from each other in nature. However, they share with the common problems that: their performance is highly dependent on the density of the observation network (e.g. DWD, 2000) and the complexity of the underlying terrain has directly impacts on the precipitation (e.g. Kyriakidis et al., 2001; Roe et al., 2005). Especially in regions with complex terrain and very large altitude gradients, very few or even no meteorological stations providing reliable rainfall measurements can be available.

In recent years, with the development of radar technology, precipitation fields from weather radar are widely used and supposed to be a good supplement with the ability to cover areas with complex terrain in fine spatio-temporal resolutions so that the patterns of rainfall are realistically reproduced (e.g. DWD, 2000; Vogl et al., 2012). However, multiple sources of errors exist in the precipitation fields derived from radar reflectivities, such as the empirical reflectivity/rainfall (Z/R) relationship, errors induced by the radar measurement itself such as backscatter or shadowing effects (e.g. Joss and Lee, 1995; Vogl et al., 2012) and etc.. When radar fields are used as the inputs to drive meteorological or hydrological models, these errors have to be taken into account carefully (e.g. Cole and Moore, 2008; Singh et al., 1997) as they will directly be propagated and can increase the uncertainty of the predicted variables.

As a result, it is straightforward and logical to join the advantages of both measurement devices, the point accurate rain gauge and spatial superior radar fields. In the past, many methods have been developed to jointly use both data sources combining their advantages, from simple to very sophisticated ones. They can be roughly classified into following different classes: gauge adjusted radar, geo-statistical approaches and the other methods (especially for Copula-based approaches in this study).

The first group was trying to reduce the uncertainties in the radar derived precipitation fields by using information from gauge. This can be done by improving the radar Z/R relationship by using rain gauge measurements (e.g. Brandes et al., 1975; Marx et al., 2007; DWD, 2000 and 2001) to correct for errors in the radar absolute values. However, the Z/R relationship is strictly dependent on the rainfall type and can have high spatial variations even in one event because of the topography and etc.. Another way is the multiple adjustment (e.g. Collier et al., 1986; Moore et al., 1994b; DWD, 1998) by calculating gauge based correction factor to modify the radar fields, trying to reduce the mean field bias or the other means (e.g. Erxleben et al., 2002). Although this kind of approach can provide satisfying results in certain specific test regions, the assumption behind is simple. The applications of these methods are all dependent on the density of rain gauge network and therefore their performances are limited in the data sparse regions.

The second way is the geo-statistical methods developed to account for the different sampling properties of radar and rain gauge, especially for the Kriging based approaches. The so called Co-Kriging was developed to combine radar and gauge data

(e.g. Krajewski et al., 1987), with the assumptions that the rainfall field is second order stationary, random errors for the rain gauges, uncorrelated in time with zero mean and etc.. Alternatively, in the work e.g. done by Seo et al. (1990a and 1990b), Universal Kriging was used to combine radar and gauge data. Furthermore, by using both gauge data and digital elevation model, Goovaerts et al. (1999) developed an extensive cross-validation comparison of different rainfall interpolation techniques. However, sophisticated Kriging with the additional complexity did not pay off in the form of better results compared to simpler Kriging techniques (e.g. Ehret et al., 2002). Furthermore, the crucial problem is that the distribution of precipitation is assumed to be Gaussian or alternative versions of Gaussian in those traditional approaches.

However, the hydro-meteorological variables are usually non-Gaussian (e.g. AghaKouchak et al., 2010b) and also require bi or multivariate analyses as well as conditional probability distributions of variables (e.g. Genest and Favre, 2007) The Copula-based approach, with the ability to capture non-linear behaviour, is not limited to the Gaussian distribution (e.g. Nelson et al., 1999). Over the past decades, there is a remarkable increase in applications of Copulas in hydro-meteorology. Alternatively, the Copulas were used to describe the complex spatio-temporal dependence structure and assess for non-linear behaviour (e.g. Genest and Favre, 2007; Dupois et al., 2007). The Copula based methods are advantageous in many respects and have already been used successfully in risk assessment (e.g. Frees and Valdez, 1998).

Roughly about 10 years ago, the Copulas started its application in the field of hydro-meteorology. Recently a remarkable increase in the Copulas related publications can be found for multivariate frequency analyses, geo-statistical interpolation and multivariate extreme value analyses (e.g. Michele and Salvadori, 2003; Dupois et al., 2007; Bárdossy et al., 2006; Genest and Favre, 2007; Renard and Lang, 2007; Schoelzel and Friederichs, 2008; Bárdossy and Li, 2008; Zhang and Singh, 2008). Considering about precipitation field, in the work done by Michele and Salvadori (2003), Copulas were used to model intensity-duration of rainfall events. Favre et al. (2004) developed Copula-based approach for multivariate hydrological frequency analysis, and then, Zhang and Singh (2008) utilised the Archimedean Copulas to perform the bivariate rainfall frequency analyses. Renard and Lang (2007) investigated the usefulness of the Gaussian Copula in extreme value analyses and Kuhn et al. (2007) tried to describe spatial and temporal dependence of weekly precipitation extremes by using the Copula-based approach. In

the work by Serinaldi et al. (2008), the author suggested a Copula-based mixed model for modelling the dependence structure and the corresponding marginal distributions, which is based on the properties of the non parametric Kendall's rank correlation and the upper tail dependence coefficient calculated from the dependence structure between pair wise observations from different rain gauges. Recently, Copula-based models were developed for the purpose to estimate the point scale or spatial distribution of errors in the radar derived precipitation fields (e.g. Villarini et al., 2008; AghaKouchak et al., 2010a and 2010b).

However, in most of these studies, the dependence structure is calculated between two variates in the bivariate framework, only with few examples of multivariate applications. Bárdossy developed a new Copula-based geo-statistical interpolation method (e.g. Bárdossy et al., 2006; Bárdossy and Li, 2008; Bárdossy and Pegram, 2009), especially the new Copula family named v-transformed Copulas. In their work, the multivariate Copulas were used to describe the spatial variability of environmental variables, e.g. groundwater quality parameters. Then a methodology was also proposed to perform the spatially interpolation for these quantities and the gauge observed daily precipitation.

Recently, the new remote sensing techniques, such as dual microwave links (e.g. Holt et al., 2003; Rahimi et al., 2004; Minda et al., 2005; Leijnse et al., 2007a) and wireless communication networks (e.g. Messer et al., 2006), show to be good complementaries to gauge and radar. The estimation of spatial rainfall fields based on the microwave link attenuation was proposed in the work as e.g. done by Giuli et al. (1991, 1999 and 2007) and Zinevich et al. (2008), also with the effort to adjust radar field by using MW-link derived rainfall (e.g. Cummings and Upton, 1996). Even though some uncertainties are still associated with their application to derive precipitation information, the widely applications of MW-link attenuation are very promising, not only due to the network of observations especially in the mountain areas but also because of very low costs comparing to gauge and radar, nearly free for commercial wireless communication networks (e.g. Messer et al., 2006).

1.3 Objectives

This study aims to develop stochastic approaches for estimating precipitation fields through assimilating data from various rainfall observation devices. Especially in the region with complex terrain, there are limited numbers of observation tools such as rain gauges. Usually, large differences are found here between precipitation fields derived from different devices. So, it is necessary and also the main objective to estimate the precipitation fields combing the advantages from different measurement tools.

Considering the point accurate rain gauge and spatial priority radar, the Copula-based analysis is chosen and capable to combine the advantages from both sides by investigating the dependence structures between them. As a result, the overall objectives of this study are:

1. Revealing the spatial distribution of dependence structures between rain gauge and radar.
2. Deriving the spatial distribution of marginal distribution for rain gauges.
3. Investigating the sensitivity of ARMA-GARCH transformation to the simulated precipitation fields.
4. Improving the estimated precipitation fields by including the precipitation information further from MW-Links.

1.4 Structure of the Dissertation

With respect to the scope of this study, the thesis is divided into 11 chapters. After the introduction and innovation, in the 3rd chapter, an overview on Copulas theory is presented, as well as the Copula-based simulation techniques and validation methods. In Chapter 4, traditional data sources (radar and gauge) and their data pre-processing are briefly introduced, including the demonstration why the gauge can be used to determine the wet/dry periods and the assumption of independent and identically distributed (*i.i.d.*). In the 5th chapter, the physical background of MW-link is introduced, followed by the comparisons between MW-link derive precipitation and gauge/radar. In Chapter 6, the point wise assimilation to combine precipitation information from rain gauge and radar is introduced and the impacts from ARMA-GARCH transformation

are investigated. Afterwards, two Copula-based spatial data assimilation approaches are proposed, which are the Copula parameter map based method in Chapter 7 and interpolated Copula parameter field based algorithm in Chapter 8. The impacts from ARMA-GARCH transformation are also investigated in the spatial case. In the 9th chapter, several spatial distributions of rain gauge marginal distributions are developed, and then the simulation results from combinations of different spatial dependence structures and gauge marginal distributions are presented and compared. The precipitation derived from MW-links is integrated to further improve the estimated rainfall fields in the 10th chapter. The final chapter is devoted to summary and conclusions and recommendations for further research.

2

Innovation

Innovative work of this thesis includes:

1. Applying Copula analysis on the hourly precipitation observations, beyond the limitation of Gaussian assumption implied in the traditional geo-statistical approaches.

2. Establishing the Copula parameter map and interpolated Copula parameter field based approaches to simulate precipitation fields of pseudo observation assimilating observations from gauge and radar, with a superior performance and also very low computational costs.

3. Developing the spatial distribution of marginal distributions for rain gauges driven by the temperature and altitude.

4. Performing ARMA-GARCH transformation to remove autocorrelation and heteroskedasticity and also investigating the sensitivity of ARMA-GARCH transformation to the final precipitation fields.

5. Including MW-link attenuation as an independent and complementary precipitation observation tool to further constrain and improve the precipitation fields derived from gauge and radar.

3

Review of Copula Theory

3.1 Introduction

The definition of Copulas first appeared in 1959 in the work of Sklar et al. (1959), which are used to model the dependence structure between two or more variables. The term of Copulas originally derived from Copulare, a Latin word which means to join, connect or link. The use of Copulas makes it possible to calculate bivariate or multivariate distributions of random variables with great flexibility. Traditional bivariate or multivariate distribution families, such as bivariate normal, log-normal and gamma, are built with a number of model parameters that describe the behaviour of each random variable as well as the joint probability distribution itself (e.g. AghaKouchak et al., 2010b). The main disadvantage of those approaches is that modelling the dependence structure between variables is not independent of the choice of the marginal distributions (e.g. Genest and Favre, 2007). However, the Copula based approaches allow to avoid this restriction and also not limited by the Gaussian assumption. Furthermore, by using the Copulas, the dependence structure keeps the same after performing any monotonic increasing transformations on the variables. As already described in section 1.2, the application of Copulas, in multivariate simulation, extreme value analysis and modelling of the dependence structure have become popular in the field of hydrology and meteorology.

In section 3.2, the theoretical background of Copula theory is given. Afterwards the Copula families used in this study are described in section 3.3, followed by the introduction of marginal distribution in section 3.4. Then the advantages of Copulas

are summarized in section 3.5. Finally, the algorithm for Copula-based simulation of pseudo observations is explained step by step in section 3.6.

3.2 Copula Theory

The Copula theory is briefly introduced in this section. First, Sklar's theorem is explained, followed by the bounds and properties for Copulas. The method to build empirical Copula is described and several important concepts and definitions are given, such as Kendall's τ Spearman's ρ and upper/lower tail dependence coefficients. More detailed information about Copula theory can be found e.g. in Joe et al. (1997); Frees and Valdez (1998); Nelson et al. (1999); Salvadori et al. (2007).

3.2.1 Sklar's Theorem

Considering a vector random variable, $X = [X_1, ..., X_n]'$, with joint distribution $F(x_1, ..., x_n)$ and marginal distributions $F_{X_i}(x_i)$, according to the theorem developed by Sklar et al. (1959), the multivariate distribution F can be expressed in terms of a Copula C and the corresponding marginal distributions:

$$F(x_1, ..., x_n) = C(F_{X_1}(x_1), \ldots, F_{X_n}(x_n)), \tag{3.1}$$

The n-dimensional Copulas are functions defined as listed in Eq. (3.2), linking univariate distribution functions together to form a multivariate distribution function.

$$C : [0, 1]^n \to [0, 1]. \tag{3.2}$$

As for general distribution functions, the density of a Copula $C(u, v)$, here $u = F_{X_1}(x_1)$ and $v = F_{X_2}(x_2)$, is calculated as

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v}. \tag{3.3}$$

Therefore, the probability density (PDF) of the multivariate distribution $f(x_1, ..., x_n)$, can be expressed in terms of a Copula PDF c and PDF of marginal distributions $f_{X_i}(x_i)$. So the Copula PDF c is often called the dependence function.

$$f(x_1, ..., x_n) = c(F_{X_1}(x_1), \ldots, F_{X_n}(x_n)) f_{X_1}(x_1) \ldots f_{X_n}(x_n) \tag{3.4}$$

With the ability to link multivariate distributions to their marginals, Copulas also allow to merge the dependence structure from the marginal distributions to form their joint multivariate distribution. The Copula function is unique when the marginals are steady functions. As the Copula is only a reflection of the dependency structure itself, their construction is reduced to the study of the relationship between the correlated variables, giving freedom for the choice of the univariate marginal distributions. This is the main advantage of Copula-based approaches.

3.2.2 The Frechet-Hoeffding Bounds for Copulas

The Frechet-Hoeffding bounds describe the upper and lower bounds for every Copula with u and v elements (the bivariate cases are taken an example in this Chapter and can be easily extended to multivariate cases). These bounds are described by

$$max(u + v - 1, 0) \leq C(u, v) \leq min(u, v) \tag{3.5}$$

where, $W(u, v) = max(u + v - 1, 0)$ is the lower bound and the upper bound is given by $M(u, v) = min(u, v)$. It should be noted that $W(u, v)$ and $M(u, v)$ are Copulas themselves.

3.2.3 Properties of Copulas

The Copula approach allows to account for the fact that the dependence structure between two variates $(u = F_X(x), v = F_Y(y))$ is more complex than it can be modelled by the multivariate normal distribution or ordinary dependency measures such as e.g. the Pearson correlation coefficient. Another important property of Copula functions is the fact that they are invariant under increasing monotonic transformations. In practice, it means that data may be transformed (e.g. by taking the logarithm or detrending) without changing its Copula. The other general properties of Copulas as e.g. given by Genest and Rivest (1993) are:

1. In the bivariate case, a Copula is defined as a function C from $[0, 1]^2$ to $[0, 1]$ so that $\forall u, v \in [0, 1]$:

$$C(u, 0) = 0 = C(0, v) \tag{3.6}$$

$$C(u, 1) = u \quad \text{and} \quad C(1, v) = v \tag{3.7}$$

2. A Copula is 2-increasing means that $\forall u_1, u_2, v_1, v_2 \in [0,1]$ with $u_1 \leq u_2$ and $v_1 \leq v_2$ holds

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) - C(u_1, v_1) \geq 0. \tag{3.8}$$

3. A Copula is continuous. Thereby satisfying the stronger Lipschitz condition.

$$|C(u_2, v_2) - C(u_2, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|. \tag{3.9}$$

4. A Copula has a survival Copula, $\overline{C}$, given by

$$\overline{C}(u, v) = 1 - u - v + C(1 - u, 1 - v) \tag{3.10}$$

And the joint survival function, $\overline{C}$, for two uniform random variables $(0, 1)$, is:

$$\overline{C}(u, v) = P[U > u, V > v] = 1 - u - v + C(u, v) = \overline{C}(1 - u, 1 - v) \tag{3.11}$$

Therefore,

$$\overline{C}(u, v) = C(1 - u, 1 - v) \tag{3.12}$$

3.2.4 Empirical Copulas

The empirical Copula $C_n(u, v)$, which is defined on the rank space (or the corresponding CDF values), is an estimator for the unknown theoretical Copula distribution $C_\theta(u, v)$ associated with the pair (X, Y) having a set of parameters θ (e.g. Genest and Rivest, 1993; Laux et al., 2011):

$$C_n(u, v) = 1/n \sum_{i=1}^{n} 1\left(\frac{r_i}{n+1} \leqslant u, \frac{s_i}{n+1} \leqslant v\right) \tag{3.13}$$

where $(r_1, s_1), \ldots, (r_n, s_n)$ denote the pairs of ranks of the data $(x_1, y_1), \ldots (x_n, y_n)$, and $1(\ldots)$ is the indicator function.

3.2.5 Kendall's τ and Spearman's ρ

There is a functional relationship between the classical dependence parameters such as Kendall's τ and Spearman's ρ (e.g. Genest and Rivest, 1993; Laux et al., 2011).

$$\rho = 12 \int_{[0,1]^2} u\, v\, dC_\theta(u, v) - 3 = 12 \int_{[0,1]^2} C_\theta(u, v)\, dudv - 3 \tag{3.14}$$

and

$$\tau = 4 \int_{[0,1]^2} C_\theta(u, v)\, dC_\theta(u, v) - 1 \tag{3.15}$$

3.2.6 Upper/Lower Tail Dependence

Tail dependence relates the amount of dependence in the upper-right-quadrant tail or in the lower-left-quadrant tail for a bivariate distribution. The upper and lower tail dependence parameters of the random vector (X, Y) with Copula C, can be defined in the following way (e.g. Joe et al., 1997):

$$\lambda_{up} \equiv \lim_{u \to 1^-} P(Y > F_Y^{-1}(u) | X > F_X^{-1}(u)) = \lim_{u \to 1^-} \frac{1 - 2u + C(u,u)}{1-u} \tag{3.16}$$

and

$$\lambda_{low} \equiv \lim_{u \to 0^+} P(Y \leq F_Y^{-1}(u) | X \leq F_X^{-1}(u)) = \lim_{u \to 0^+} \frac{C(u,u)}{u} \tag{3.17}$$

The upper tail dependence expresses the probability occurrence of positive large values (outliers) at multiple locations jointly while the lower tail dependence expresses the the probability occurrence of positive small values.

3.3 Copula Families

The two most commonly used Copula families are the Elliptical and the Archimedean Copula families. These Copulas are called Elliptical because they can be constructed from elliptical distributions. Since this Copula family closely related to the multivariate normal distribution, it provides important examples of multivariate distributions. However, the Elliptical Copula shows a symmetrical upper/lower tail dependence structure so that the application of this Copula family is limited.

Archimedian Copulas are defined as following, let $\varphi : [0,1] \to [0,\infty]$ a steady, strict monotonic function with $\varphi(1) = 0$ and let $\varphi^{[-1]} : [0,\infty] \to [0,1]$ the pseudo-inverse of φ (e.g. Nelson et al., 1999):

$$\varphi^{[-1]} := \begin{cases} \varphi^{-1}(t) & if \quad 0 \leq t \leq \varphi(0), \\ 0 & else \end{cases} \tag{3.18}$$

then the function

$$\begin{array}{rcl} C : [0,1]^2 & \to & [0,1] \\ (u,v) & \mapsto & \varphi^{[-1]}(\varphi(u) + \varphi(v)) \end{array} \tag{3.19}$$

defines a Copula only if φ is convex and φ is called the generator of the Archimedian Copula C.

3.3.1 Gaussian Copula

The Gaussian Copula is derived from the multivariate normal distribution φ_R. The CDF of the Gaussian Copula is defined as:

$$C(u, v) = \varphi_R(\phi^{-1}(u), \phi^{-1}(v)) \tag{3.20}$$

Where R is the linear correlation matrix and ϕ^{-1} is the inverse of the univariate standard normal distribution function. Note that the equation above is for the bivariate case; for the multivariate case, equation can be written as

$$C(u_1, \cdots, u_n) = \varphi_R(\phi^{-1}(u_1), \cdots, \phi^{-1}(u_n)) \tag{3.21}$$

If the marginal distributions $u_1 = F_1(x_1), \cdots, u_n = F_n(x_n)$ are normal, then the random vector $(x_1, , x_n)$ has a multivariate normal distribution.

3.3.2 Student-T Copula

The Student-T Copula is derived from the multivariate t-distribution. The CDF of the Student-T Copula can is defined as:

$$C_\Theta(u, v) = t_{\nu,\Sigma}(t_\nu^{-1}(u), t_\nu^{-1}(v)) \tag{3.22}$$

$\Theta = \{(\nu, \Sigma) : \nu \in (1, \infty), \Sigma \in \Re^{m \times m}\}$, t_ν is a univariate t distribution with ν degrees of freedom and $t_{\nu,\Sigma}$ is the multivariate t distribution with a correlation matrix Σ with ν degrees of freedom

3.3.3 Gumbel Copula

The CDF of the Gumbel Copula is defined as

$$C_\theta(u, v) = e^{-\left((-\ln(u)^\theta) + (-\ln(v)^\theta)\right)^{\frac{1}{\theta}}} \tag{3.23}$$

with $\theta > 1$. It is an Archimedean Copula with the generator $\varphi(t) = (-\ln(t))^\theta$. This Copula is usually used for asymmetrical tail dependence structure (e.g. Nelson et al., 1999).

3.3.4 Clayton Copula

The Clayton Copula is defined as

$$C_\theta(u,v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}} \tag{3.24}$$

where $\theta > 0$. It is also an Archimedean Copula with the generator $\varphi(t) = \frac{1}{\theta}\left(t^{-\theta} - 1\right)$. The Clayton Copula is known to be asymmetrical and has a lower tail dependency

3.3.5 Frank Copula

The Frank Copula is defined as

$$C_\theta(u,v) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right) \tag{3.25}$$

with $\theta > 0$. It is also an Archimedean Copula with the generator $\varphi(t) = -\ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)$ Note that as $\theta \to 0$ the dependence becomes maximal and for $\theta = 1$ independence is achieved. This Copula is known to be symmetrical.

3.4 Marginal Distributions

Generally, e.g. for a bivariate joint distribution of (X, Y), the marginal distribution of X is calculated by summing or integrating the joint probability distribution over Y, and the same for Y. For discrete random variables, the marginal probability mass function can be written as $Pr(X = x)$. This is

$$Pr(X = x) = \sum_y Pr(X = x, Y = y) = \sum_y Pr(X = x|Y = y)Pr(Y = y) \tag{3.26}$$

where $Pr(X = x, Y = y)$ is the joint distribution of X and Y, while $Pr(X = x|Y = y)$ is the conditional distribution of X given Y.

Similarly for continuous random variables, the marginal probability density function can be written as $p_X(x)$. This is

$$p_X(x) = \int_y p_{X,Y}(x,y)dy = \int_y p_{X|Y}(x|y)p_Y(y)dy \tag{3.27}$$

where $p_{x,y}(x,y)$ gives the joint distribution of X and Y, while $p_{X|Y}(x|y)$ gives the conditional distribution for X given Y. Again, the variable Y has been marginalized out.

In this study, 4 different theoretical distribution functions are tested to find the best fit and their formulas are listed below.

1. The Normal distribution with mean μ and standard deviation σ

$$f(x) := \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)} \tag{3.28}$$

2. The Exponential distribution with parameter $\lambda \in \mathbb{R}_{>0}$

$$f_\lambda(x) := \begin{cases} \lambda e^{-\lambda x} & if \quad x \geq 0, \\ 0 & else \end{cases} \tag{3.29}$$

3. The Weibull distribution with $\alpha > 0$, $\beta > 0$ and if $\beta = 1$, the Weibull reduces to Exponential.

$$f(x) := \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} \tag{3.30}$$

4. The Gamma distribution with $b, p \in \mathbb{R}$ and $b > 0$, $p > 0$, where $\Gamma(p)$ denotes the value of the Gamma function at p

$$f_\lambda(x) := \begin{cases} \frac{b^p}{\Gamma(p)} x^{p-1} e^{-bx} & if \quad x \geq 0, \\ 0 & else \end{cases} \tag{3.31}$$

All the marginal distributions of the positive rainfall amounts were estimated by maximum likelihood for each gauges, radar or MW-links at each location/grid. The distribution was chosen with the minimum value of the Akaike Information Criterion (AIC) value, which is a measure of the relative Goodness of Fit of a statistical model developed by Akaike et al. (1974). The Bayesian information criterion (BIC) is an updated version of AIC considering the number of data. The AIC and BIC are defined as:

$$AIC = 2k - 2\ln(L) \tag{3.32}$$

and

$$BIC = k\ln(n) - 2\ln(L) \tag{3.33}$$

where k is the number of parameters in the statistical model, n is the number of data in x and L is the maximized value of the likelihood function for the estimated model.

3.5 ARMA-GARCH Transformation

The algorithm for ARMA-GARCH transformation taken from Laux et al. (2011) and the detail information can be e.g. found in Engle et al. (1982) and Bollerlslev et al. (1986).

This section describes the theory of the ARMA-GARCH composite model to produce independent and identically distributed *i.i.d.* residuals. An ARMA model is used to compensate for autocorrelation, and a GARCH model to compensate for the heteroskedasticity.

The GARCH − Generalized Autoregressive Conditional Heteroskedasticity is a time series modelling technique that includes past variances for predicting present or future variances. A univariate model of an observed time series y_t can be written as

$$y_t = E(y_t|\Omega_{t-1}) + \varepsilon_t \tag{3.34}$$

In this equation, the term $E(|)$ denotes the conditional expectation operator, Ω_{t-1} the information set at time $t-1$, and ε_t the innovations at time t. Bollerlslev et al. (1986) developed GARCH as a generalization of the ARCH volatility modelling technique (Engle et al., 1982). The distribution of the residuals, conditional on the time t, is given by

$$Var_{t-1}(y_t) = E_{t-1}(\varepsilon_t^2) = \sigma_t^2 \tag{3.35}$$

where

$$\sigma_t^2 = \kappa + \sum_{i=1,P} G_i \sigma_{t-i}^2 + \sum_{i=1,Q} A_i \varepsilon_{t-j}^2 \tag{3.36}$$

where κ is a constant, and σ_t^2 is the prediction of the variance, given the past sequence of variance predictions σ_{t-i}^2, and past realizations of the variance itself ε_{t-j}^2. When $P = 0$, the GARCH $(0, Q)$ model becomes the original ARCH(Q) model introduced by Engle et al. (1982). This equation mimics the variance clustering of the variable (i.e. precipitation). The lag lengths P and Q and the coefficients G_i and A_j determine the degree of persistence.

A common assumption is that the innovations are serially independent, however, GARCH(P, Q) innovations, ε_t, are modelled as

$$\varepsilon_t = \sigma_t Z_t \tag{3.37}$$

t is the conditional standard deviation given by the square root of Eq. (3.35), and Z_t is the standardized *i.i.d.*. random draw from some specified probability distribution. Usually, a Gaussian distribution is assumed such that $\varepsilon \sim N(0, \sigma_t^2)$. Reflecting this, Eq. (3.36) illustrates that a GARCH innovations process ε_t simply rescales an *i.i.d* process Z_t such that the conditional standard deviation incorporates the serial dependence of Eq. (3.35).

3.6 Copula-Based Analysis and Simulation

In this section, the Copula-based analysis and simulation techniques are introduced. First is the Goodness of Fit (GoF) test to choose the best theoretical Copula function to model the dependence structure derived from the empirical Copulas. In the second section, the Copula-based simulation techniques are described in detail. Finally, the validation measures are presented, which are used in this study to check the quality of the simulations.

3.6.1 Goodness of Fit Tests

Goodness of Fit tests for Copulas are applied by comparing the empirical Copula C_n with the parametric estimate of a theoretical Copula model C_θ derived under the null hypothesis. There are different GoF tests (e.g. Genest and Remillard, 2008; Genest et al., 2009). One of the tests used in this study is based on the Cramér-von Mises statistic (e.g. Genest and Favre, 2007):

$$S_n = n \sum_{i=1}^{n} \{C_\theta(u_i, v_i) - C_n(u_i, v_i)\}^2. \tag{3.38}$$

As the definition of S_n involves the theoretical Copula function, the distribution of this statistic depends on the unknown value of θ under the null hypothesis that C is from the class C_θ (e.g. Grégoire et al., 2008). Therefore, the approximate p-values for the test statistic are obtained using a parametric bootstrap (e.g. Genest and Remillard, 2008; Genest et al., 2009). However, for the precipitation data used in this study, this GoF test is not sensitive enough so that another GoF test are further applied.

The second GoF test is the non-parametric method called K-function. For an Archimedean Copula with the generator φ, the K-function is defined as:

$$K = t - \frac{\varphi(t)}{\varphi^{'}(t)} \tag{3.39}$$

Where t is the value of empirical CDF, thus with n number of data (length of data), $t = i/n, i = 1, 2, ..., n$ The non-parametric estimate for the K-function is given by:

$$K^{'} = \frac{1}{n}\sum_{j=1}^{n}\vartheta_j \tag{3.40}$$

$$\vartheta_j = \frac{1}{n-1}\sum_{i}^{n}\{(x,y) : x < x_i, y < y_i\} \tag{3.41}$$

The best fitting Archimedean Copula is determined based on the L between the empirical and theoretical K function values. The smallest L means the best fit.

$$L(K, K^{'}) = \sum_{i}^{n}(K - K^{'})^2 \tag{3.42}$$

3.6.2 Simulating From Copula Distributions

This section is also taken from Laux et al. (2011) and the brief introduction is given in the following.

In practice, for bivariate case, the following 4 steps are applied to model the dependence structure:

1. The data (x_i, y_i) is transformed to the rank space (r_i, s_i) with $i = 1, \ldots, n$ denoting the length of the dataset.

2. The empirical Copula $C_n(u, v)$ is calculated using the ranks (r_i, s_i).

3. The Copula parameters are estimated using different types of theoretical Copula functions (maximum likelihood estimation).

4. GoF tests are carried out to choose the appropriate Copula family and its parameter

After the estimation of the Copula-based joint distribution - $F_X(x)$, $F_Y(y)$ and $C_\theta(u, v)$ - conditional random samples from this distribution are generated through Monte Carlo simulations. The simulation is based on conditional probabilities of the form:

$$P(U \leq u|V = v) = \frac{\partial}{\partial v} C(u, v) \tag{3.43}$$

$$P(V \leq v|U = u) = \frac{\partial}{\partial u} C(u, v) \tag{3.44}$$

The concept for generating pseudo-observation is as follows: a pair of variates (u, v) in the rank space with Copula $C_\theta(u, v)$ needs to be generated which finally can be transformed into (x, y) in the data space, using the probability integral transformation

$$U = F_X(x) \quad \Leftrightarrow \quad X = F_X^{-1}(U = u) \tag{3.45}$$

$$V = F_Y(y) \quad \Leftrightarrow \quad Y = F_Y^{-1}(V = v). \tag{3.46}$$

With the estimated theoretical Copula function - $c_\theta(u, v)$ - and the marginal distributions - $F_X(x)$, $F_Y(y)$ -, conditional random samples are generated using the conditional probability density function. The algorithm for simulating pseudo observation from one of the variables is as follows:

1. Computation $u = F_X(x)$, where x denotes one value of the radar measured rainfall and $F_X(x)$ is the marginal distribution of the variate X.

2. Generation of random samples for the variate v^* from the conditional PDF $c_{V|U}(v|u) = c_u(v)$ and calculation of $v = c_u^{-1}(v^*)$, where c_u^{-1} denotes the generalized inverse of c_u (Nelsen, 1999).

3. Calculation of the corresponding y-values using the probability integral transformation $F_Y^{-1}(v) = y$.

The final result for y is a sample of pseudo observations which lies in the original data space and can be compared with the observed gauge series. This method may be used e.g. for generating pseudo observations conditioning on model data to bias-correct the result or for downscaling GCM results (Laux et al., 2011). This thesis adopts the procedure to estimate precipitation fields by integrating data from radar, gauge and MW-link, in both point wise and spatial scale.

3.6.3 Validation Measures

In order to test the efficiency of the Copula-based approaches, simulated fields of pseudo-observations are compared to the original radar measurements, improved radar fields (e.g. Radolan), the rain gauge observations in the study area. The quantitative validation is performed by using the point wise cross-validation. Table 3.1 shows the different efficiency criteria used in this study, which are Pearson's Correlation (r), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Nash-Sutcliffe Efficiency (NSE).

Table 3.1: Validation measures used in this study, (o_i define the value of the observations and m_i the value of the model at time step $i = 1, \ldots, n$).

Validation measure	Formula	Range	Perfect Fit
r	$r = \frac{\sum_{i=1}^{n}(o_i-\bar{o_i})(m_i-\bar{m_i})}{\sqrt{\sum_{i=1}^{n}(o_i-\bar{o_i})^2}\sqrt{\sum_{i=1}^{n}(m_i-\bar{m_i})^2}}$	$[-1, 1]$	$\lvert r\rvert = 1$
RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - m_i)^2}$	$[0, \infty[$	RMSE=0
MAE	$\frac{1}{n}\sum_{i=1}^{n}\lvert o_i - m_i\rvert$	$[0, \infty[$	MAE=0
NSE	$1 - \frac{\sum_{i=1}^{n}(o_i-m_i)^2}{\sum_{i=1}^{n}(o_i-\bar{o_i})^2}$	$]-\infty, 1]$	NSE=1

4

Study Area and Traditional Data Sources

This chapter gives an overview of the study area and traditional data sources used in this thesis. This includes information not only about the nature and characteristics of the study area, but also about the observations from radar, rain gauges and meteorology stations by German Weather Service (Deutscher Wetterdienst, DWD). In section 4.1, details about the study area are given. Afterwards, data sources from DWD are described in the following section 4.2.1 and 4.2.2, radar and gauge separately, as well as their pre-processing steps in section 4.2.3. The radar/gauge pair-wise comparisons are presented in section 4.3, proving that gauge observations can be used to determine rainy or not for each time step. Then, in section 4.4, the assumption of the *i.i.d.* is discussed. Finally, a brief summary of this chapter is given in section 4.5.

4.1 Alps and Alpine Forelands

In regions where precipitation reveals high spatio-temporal variability, such as alpine or pre-alpine terrain, it is a specific challenge to estimate realistic rainfall fields. Therefore the study area is chosen to be the southern Bavarian Alps and alpine forelands in Germany, around Garmisch-Partenkirchen and the Ammer catchment, shown in Figure 4.1, to test the model's reliability, robustness and applicability. There are large gradients in elevation across the domain with the lowest values in the north. Within the flat area of the domain, the altitude of Munich is at 519 m.a.s.l. In turn, the highest

location is Mount Zugspitze at 2962 m.a.s.l.

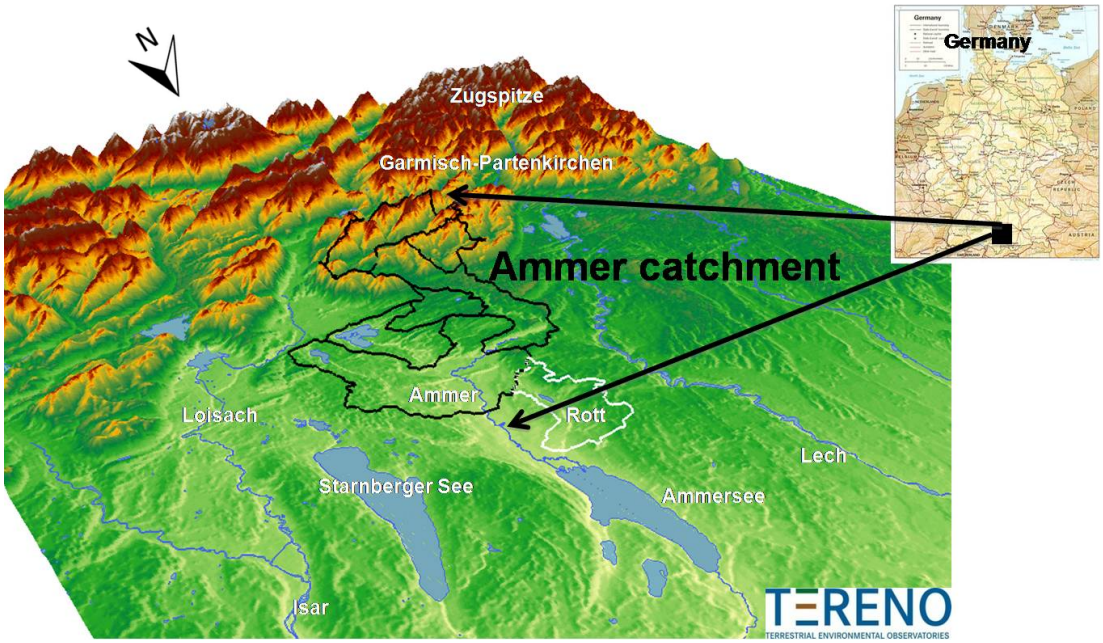

Figure 4.1: Southern Bavarian Alps and alpine forelands, Germany.

The climate in the catchment can be characterized as cool-temperate and sempiternal humid. Due to the relief, all climate variables have latitude and height dependent gradients. The mean annual temperature is around 7-8 $°C$ in the alpine forelands and 4.5 $°C$ in the southern part of the catchment. The temperature distribution shows a height-dependent gradient of around 0.6 $°C/100$ m in summer and 0.45 $°C/100$ m in winter. There are around 130 days with snow cover (snow depth $>$ 10 cm) in the southern part of this area each year.

Due to the complex orography and heterogeneity in topography, the region is characterised by big north southerly differentiations in soils, land-use, climate and also precipitation. Long term mean annual precipitation in the northern part of the catchment is around 1100 mm while the southern part with the summits of the Ammer Alps receives more than 2000 mm. Figure 4.2 shows the two exemplary distributions of mean monthly rainfall at Garmisch-Partenkirchen and Hohenpeissenberg located in the study area. Maximum precipitation is reached in June, July and August, which is one of the reasons to choose summer time as the focus of this study.

4.2 Traditional Data Sources

Both C-band weather radar and rain gauge measured hourly precipitation are provided by DWD in this region as listed in Table 4.1. Besides precipitation, hourly air temperature observations are available for specific stations.

Figure 4.3 shows an overview over the study area with the marked positions of the available gauges (black dots) and the position of the weather radar at Mount Hohenpeissenberg (red triangle). In the next sub-sections, those different kinds of traditional data sources are described in detail.

4.2.1 Radar

The radar data used in this study was taken from a C-band research weather radar operated by the Meteorological Observatory Hohenpeissenberg (MOHP), DWD. The observatory is located about 80 km south-west of Munich on top of mount Hohenpeissenberg at an altitude of 977 m.a.s.l. The radar installation covers a circular area with a radius of 256 km, producing a scan every 5 min. For the study area, a square region of 100 x 100 grid cells with 1 km x 1 km resolution centred at the radar station at Mount Hohenpeissenberg is selected. As can be seen in the bottom of Figure 4.5, the radar is in the middle of all gauges. The technical data of this C-band radar is given in Table 4.2.

Preprocessing operations such as the clutter correction of the radar reflectivities have been performed by DWD. After that, rainfall amounts are derived by using the DWD standard Z (Reflectivity in dbz) / R (Rainfall in mm/hour) relationship being

$$Z = 256R^{1.42} \tag{4.1}$$

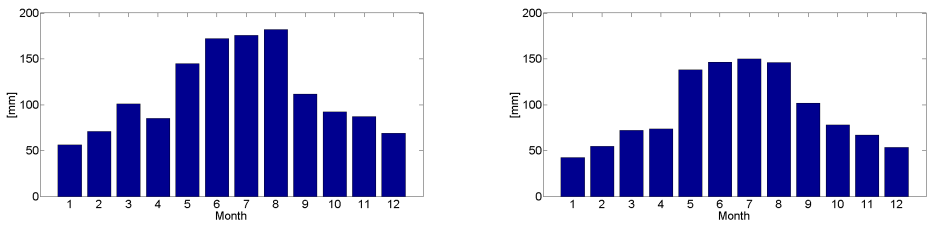

Figure 4.2: Distribution of mean monthly precipitation [mm] at Garmisch-Partenkirchen (left) and Hohenpeissenberg (right), 1996-2010.

Table 4.1: The description of available rain gauge stations and air temperature observations in this study area, both in hourly.

ID	Station Name	Altitude [m.a.s.l]	Lat [°N]	Lon [°E]	Temperature Y/N
1	Bernbeuren-Prachtsried	936	47.74	10.75	N
2	Diessen	658	47.96	11.01	N
3	Deisenhofen	585	48.04	11.58	N
4	Ettal	940	47.57	10.96	N
5	Garmisch-Partenkirchen	719	47.48	11.06	Y
6	Gilching	550	48.11	11.28	N
7	Griesen	801	47.48	10.95	N
8	Halblech	780	47.65	10.81	N
9	Hindelang	1015	47.46	10.43	N
10	Hohenpeissenberg	977	47.80	11.01	Y
11	Kaufbeuren	716	47.87	10.60	Y
12	Kochel	805	47.57	11.30	N
13	Kohlgrub, Bad	740	47.67	11.08	N
14	Kraftisried	831	47.77	10.46	N
15	Kreuth	895	47.61	11.65	N
16	Krün	873	47.50	11.28	N
17	Lenggries	737	47.59	11.55	N
18	Maisach	530	48.21	11.20	Y
19	Marktoberdorf	790	47.72	10.64	N
20	München	515	48.16	11.54	Y
21	Oberammergau	835	47.60	11.06	N
22	Oberschleissheim	484	48.24	11.55	N
23	Oy	885	47.64	10.39	Y
24	Schwangau	796	47.58	10.72	N
25	Seeg	802	47.67	10.63	N
26	Schäftlarn	557	47.98	11.47	N
27	Steingaden	761	47.76	10.86	N
28	Schwaben	538	48.20	10.73	N
29	Schlehdorf	609	47.66	11.32	N
30	Vilgertshofen	685	47.97	10.92	N
31	Wielenbach	550	47.88	11.16	Y

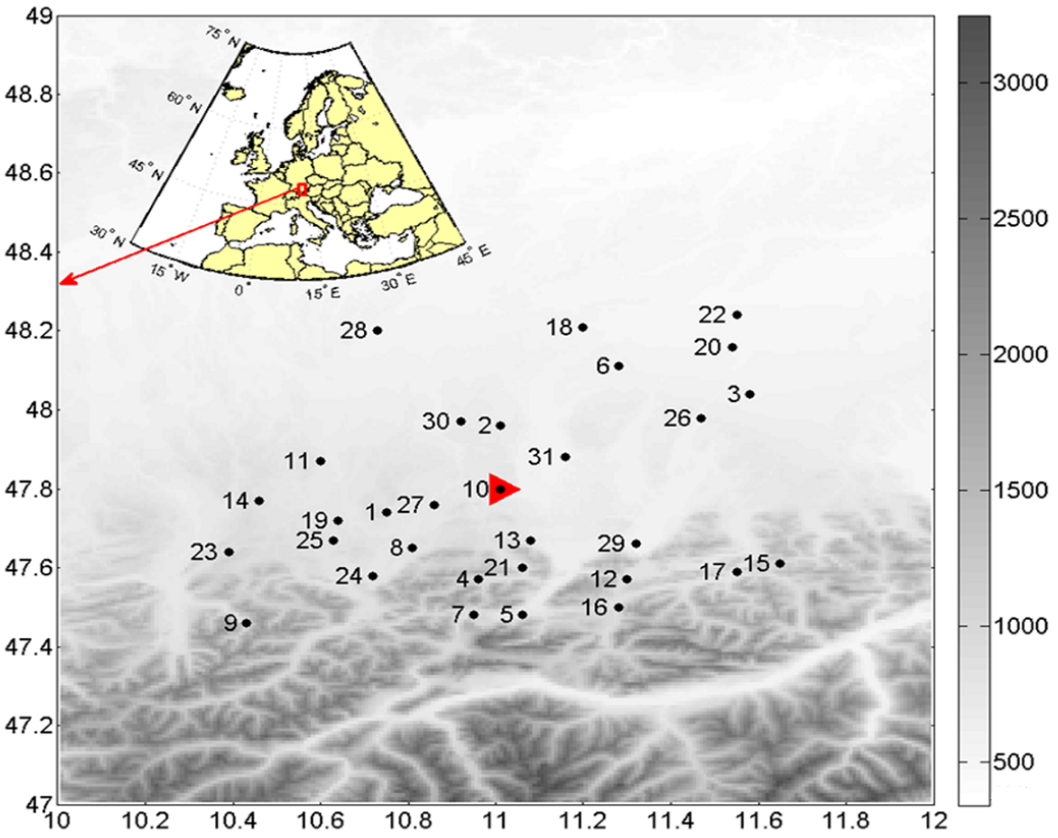

Figure 4.3: Research area showing the position of the gauges (black dots) and the weather radar (red triangle) on mount Hohenpeissenberg (Vogl et al., 2012). The names of the gauge stations can be found in Table 4.1. The unit of the color bar is in [m.a.s.l].

The data covers the time period from 2005 to 2009 as shown in the top of Figure 4.4, with some missing time steps due to the change of hardware. An exemplary precipitation field derived from radar reflectivities measured at mount Hohenpeissenberg on 14/07/2008, 13:00 is shown in the bottom of Figure 4.4. The coordinate system is centred at the position of the weather radar (white triangle). The red coloured areas mark the regions with high precipitation. The radar field is disturbed by spikes where obstacles are shading the radar beam. There may be also effects of backscatter even if the field is corrected before. However, the measured radar field gives a realistic impression of the precipitation field at that certain time step.

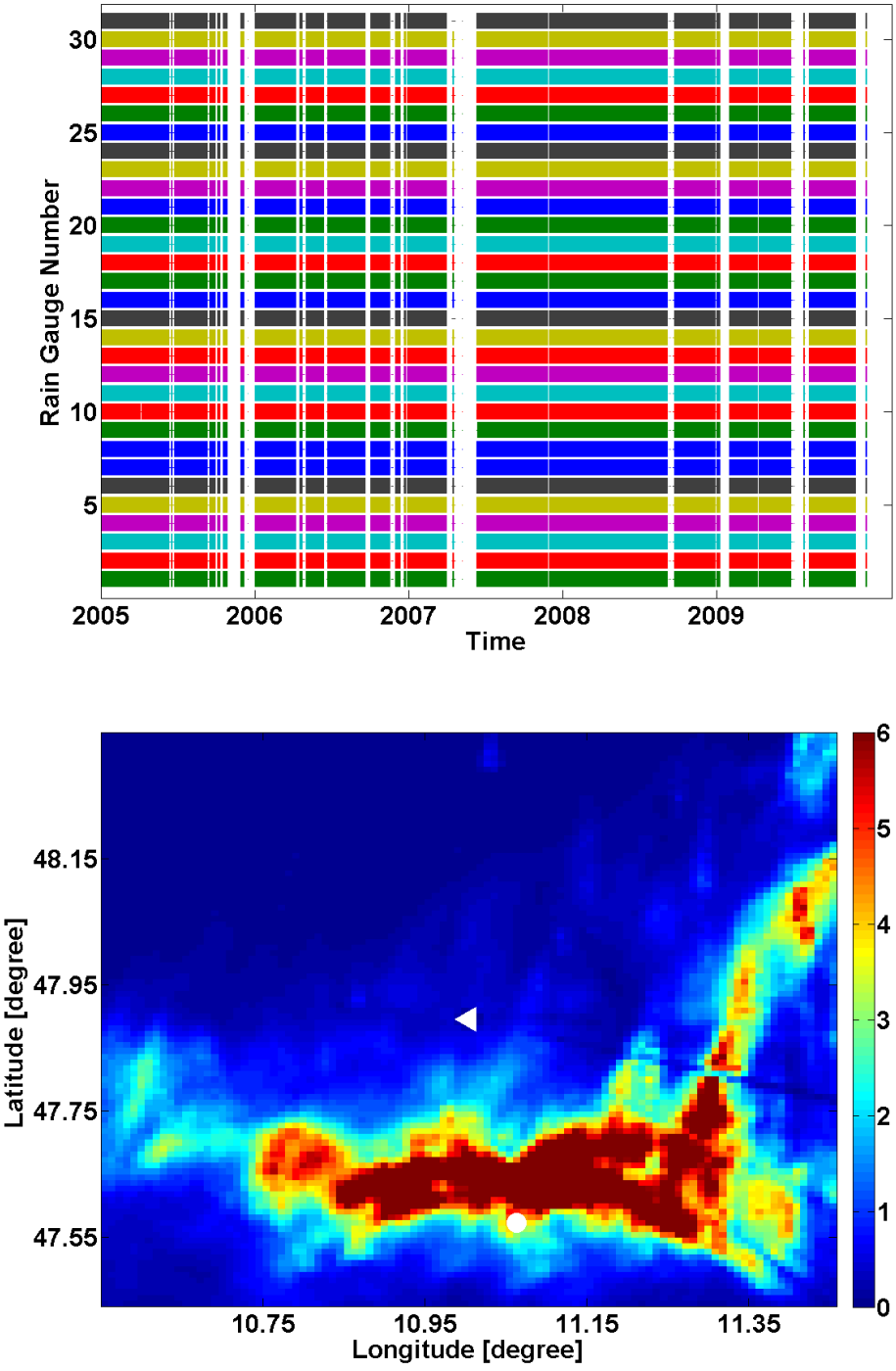

Figure 4.4: Available time period of radar measurements for the grids with gauges (top, arranged by the station ID in Table 4.3) and precipitation field derived from radar reflectivities measured at mount Hohenpeissenberg on 14/07/2008, 13:00 (bottom). The white triangle refers to radar station Hohenpeissenberg and the white cycle refers to Garmisch-Partenkirchen, the same all through the thesis. The unit of the color bar is in [mm/hour].

Table 4.2: Technical descriptions of C-band weather radar at Hohenpeissenberg

Specification	value
Antenna diameter	4.2 m
Beamwidth	0.1°
Pulse Repetition Frequency	1000 Hz
Radom diameter	6 m
Range	128 km
Receiver MDS	10^{-4} W
Spatial Resolution	1 km
Temporal Resolution	5 min
Transmitted power	200 to 250 kW
Wave length	5.3 cm

4.2.2 Gauge

Precipitation data of 31 gauges within the chosen domain is retrieved from the Web-werdis portal of DWD (http://werdis.dwd.de/) which covers the same period or even longer as the radar data (Figure 4.5). All gauges have the resolution at 60 min, 0.1 mm/h.

Note that this hourly gauge product is generated in a special way by averaging measurements between xx : 51 to xx+1 : 50 to be the gauge precipitation at hour xx, all in Coordinated Universal Time (UTC). So, the radar rainfall are also averaged in the same way to produce the hourly measurements.

General informations about the gauge stations (station name, ID, lon lat, altitude) are listed in Table 4.1 and it is also indicated whether the air temperature observation is available (Y) or not (N). Each gauge station is assigned to its corresponding grid cell in the radar domain. The gauge stations provide a network of continuous measurements with a vague homogeneous coverage, and an average distance between neighbouring stations of about 21 km. As the observation network is very sparse there is no grid cell assigned to more than one gauge station.

As can be seen in Figure 4.3 and Table 4.1, the density of the stations is reduced in higher elevations, which means that there are very limited or even no rain gauges

in the mountain areas. It is known that the altitude can have impacts on the rainfall distribution and intensity. Therefore this kind of the station distribution can lead to systematic errors in the traditional interpolation procedure, for instance the Kriging methods, especially in this study area with such a complex terrain.

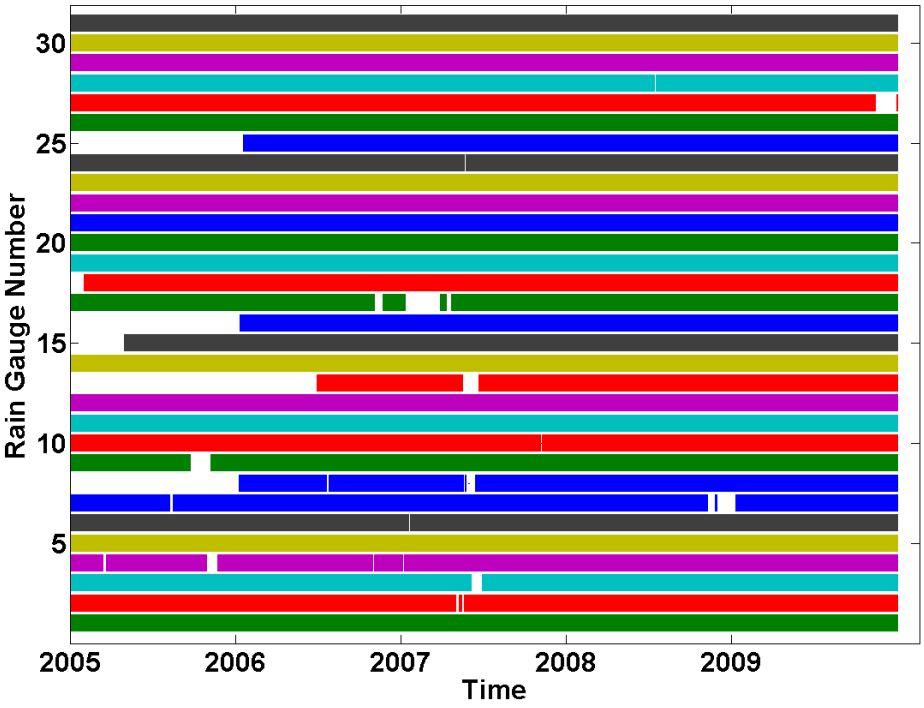

Figure 4.5: Available time period of 31 rain gauges (arranged by station ID in Table 4.1).

4.2.3 Preliminary Data Processing and Availability

For each grid with both gauge and radar precipitation measurements, the gauge/radar pairs are checked for plausibility and erroneous or significantly anomalous values are removed. This procedure consists of the following steps:

1. High gauge values are checked for plausibility by comparing with nearby gauges.

2. Radar rainfall values smaller than 0.1 mm/hour are set to NaN (Not a Number) as these measurements are considered to be erroneous.

3. The differences between neighbouring radar grid cells are calculated. Single values with absolute differences exceeding a threshold of 25 mm/hour, revealing unrealistically large gradients in the radar field, are removed (Marx et al., 2007).

4. Only the remaining positive pairs (radar and gauge) of rainfall intensities are considered for further analysis.

The remaining data pairs are divided into two subsets. The first set containing data of June, July, and August of 2006 and 2007 serves as calibration data. The second data set is containing data from the same months of 2008. Mean and standard deviation of gauge and radar (positive pairs) for the calibration period are listed in Table 4.3. Especially for station Hohenpeissenberg, the number of positive pairs is limited, only 70 pairs are available. As the gauge station is located very close to radar, the radar beam cannot capture the area above this gauge. This is the reason why mean and standard deviations of the positive pair differ significantly from those of the other stations.

4.3 Radar/Gauge Pair Wise Comparison

Due to the different measurement principles between radar and gauge, as introduced in the previous chapter, the discrepancies between radar and rain gauge can be very large (e.g. Smith et al., 1996; Wilson et al., 1979). For instance, it is likely to happen that the gauge can observe precipitation while the radar can not and vice verse. Considering one grid cell with both gauge and radar, a bivariate data set - (gauge, radar) - can be investigated. This data set can be classified to four groups according to their values lager or equal to zero. They are (0,0), (1,0), (0,1) and (1,1), where '0' represents no rain while '1' indicates the rain was observed. By computing the number of each group in the proportion of the total number of data, the probability values - p_{00}, p_{01}, p_{10} and p_{11} - for each group are listed in the Table 4.4 exemplary for the locations at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich city in different seasons, 2005 to 2008. Note that the time steps where the value have been set to be NaN are not accounted in, as they are errors or no observations.

Generally, p_{00} is always the largest value and p_{11} is the second one. It is easy to understand that the number of dry hours should be more than the number of wet hours so that the dry period is in the dominate position. However, during summer

Table 4.3: Mean and standard deviation for gauge and radar (positive pairs only) in the period of June, July, and August, 2006 to 2007.

ID	Station Name	Gauge_{mean} (mm/hour)	Gauge_{std} (mm/hour)	Radar_{mean} (mm/hour)	Radar_{std} (mm/hour)
1	Bernbeuren-Prachtsried	1.76	2.27	1.33	2.23
2	Diessen	1.51	1.96	1.15	1.74
3	Deisenhofen	1.62	2.12	1.25	1.60
4	Ettal	1.59	1.85	1.36	1.88
5	Garmisch-Partenkirchen	1.61	2.02	1.48	2.13
6	Gilching	1.56	2.25	1.25	1.75
7	Griesen	1.48	1.75	1.49	2.13
8	Halblech	1.68	2.27	1.15	1.42
9	Hindelang	1.75	2.28	1.59	2.67
10	Hohenpeissenberg	5.03	4.10	0.16	0.01
11	Kaufbeuren	1.66	2.16	1.26	1.93
12	Kochel	1.51	1.87	1.22	1.60
13	Kohlgrub, Bad	1.82	2.27	1.25	1.91
14	Kraftisried	1.70	2.25	1.30	2.12
15	Kreuth	1.72	2.13	1.43	2.11
16	Krün	1.71	2.18	1.46	1.97
17	Lenggries	1.73	2.20	1.55	2.40
18	Maisach	1.55	2.04	1.19	1.69
19	Marktoberdorf	1.77	2.30	1.32	2.08
20	München	1.53	2.21	1.43	1.86
21	Oberammergau	1.58	2.11	1.10	1.88
22	Oberschleissheim	1.55	2.13	1.24	1.81
23	Oy	1.84	2.14	1.52	2.09
24	Schwangau	1.63	2.13	1.47	2.04
25	Seeg	1.67	2.19	1.36	1.95
26	Schäftlarn	1.58	2.01	1.24	1.59
27	Steingaden	1.72	2.18	1.16	1.83
28	Schwaben	1.54	1.91	1.18	1.67
29	Schlehdorf	1.76	2.12	1.41	1.82
30	Vilgertshofen	1.66	2.16	1.21	1.71
31	Wielenbach	1.45	1.89	1.10	1.44

Table 4.4: The probabilities - p_{00}, p_{01}, p_{10} and p_{11} - for each group of the locations at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich City in different seasons from 2005 to 2008.

Season/Location	Garmisch-P.				Oberammergau			
	p_{00}	p_{01}	p_{10}	p_{11}	p_{00}	p_{01}	p_{10}	p_{11}
Spring	0.61	0.14	0.01	0.24	0.41	0.12	0.01	0.46
Summer	0.63	**0.05**	**0.01**	**0.31**	0.30	**0.10**	**0.01**	**0.59**
Autumn	0.73	0.05	0.06	0.16	0.18	0.18	0.01	0.63
Winter	0.83	0.04	0.05	0.09	0.65	0.05	0.09	0.22
Season/Location	Wielenbach				Munich			
	p_{00}	p_{01}	p_{10}	p_{11}	p_{00}	p_{01}	p_{10}	p_{11}
Spring	0.49	0.12	0.01	0.39	0.66	0.08	0.01	0.25
Summer	0.42	**0.08**	**0.01**	**0.51**	0.68	**0.07**	**0.01**	**0.25**
Autumn	0.30	0.16	0.01	0.54	0.73	0.07	0.01	0.19
Winter	0.75	0.04	0.05	0.15	0.76	0.05	0.05	0.13

and autumn at Oberammergau and Wielenbach, p_{11} is the maximum. This is due to the reason that those radar measurements between 0 and 0.1 are set to be NaN and not accounted in because those radar observations are not accurate or even error. To exclude those time steps with the unqualified radar measurements is reasonable.

Comparing the values of p_{00}, p_{01}, p_{10} and p_{11}, it can be seen from the Table 4.4 that p_{10} is always the smallest, around 0.07 for winter and less than 0.01 for other seasons for all the four listed locations. During spring, autumn and winter, the ratios of p_{10}/p_{11} and p_{01}/p_{11} are so large that those can not be neglected. This is because both radar and gauge observations are not qualified due to the impacts from snow and ice during the long period with very low temperature, especially for the mountain areas.

During summer time, both p_{01} and p_{10} are small compared to p_{11}. The mean of the ratio - p_{10}/p_{11} - is around 0.4 percent. In this case, the (1,0) case can be neglected due to this limited number in (1,0) comparing to that in (1,1) case. Meanwhile, the mean ratio - p_{01}/p_{11} - is around 17 percent, not small enough as the previous one. So, the (0,1) case has still to be considered.

Furthermore, PCA - principle component analysis (Abdi and Williams, 2010) and factor analysis are applied to analyse other climate observations such as hourly air temperature, relative humidity, mean wind speed, as well as their gradients in the cases

(0,0), (1,0), (0,1) and (1,1). The results of PCA coefficients from gauge at Garmisch-Partenkirchen, during summer from 2005 to 2008, are shown in Table 4.5.

Table 4.5: PCA coefficients for air temperature T, ΔT, relative humidity, ΔRH, mean wind speed and ΔMWS at Garmisch-Partenkirchen for (0,0), (1,0), (0,1) and (1,1) cases in different seasons from 2005 to 2008.

Coef./Season	Spring				Summer			
	(0,0)	(0,1)	(1,0)	(1,1)	(0,0)	(0,1)	(1,0)	(1,1)
T	546.6	412.9	147.3	57.0	383.9	283.9	42.5	76.8
ΔT	71.4	50.2	16.0	33.9	51.7	48.5	13.2	35.5
RH	36	21	7	15	12	13	6	8
ΔRH	1.0	0.7	0.3	0.9	0.8	0.9	0.4	0.8
MWS	0.87	0.69	0.18	0.36	0.56	0.51	0.28	0.32
ΔMWS	0.30	0.25	0.09	0.27	0.28	0.35	0.15	0.25

Coef./Season	Autumn				Winter			
	(0,0)	(0,1)	(1,0)	(1,1)	(0,0)	(0,1)	(1,0)	(1,1)
T	273.5	250.8	23.9	36.3	219.2	202.9	29.2	38.4
ΔT	46.1	43.7	16.8	29.2	33.3	28.1	10.4	23.5
RH	29	289	10	21	29	7	7	6
ΔRH	0.7	0.8	0.5	0.6	1.0	0.8	0.8	0.6
MWS	0.56	0.62	0.25	0.25	0.41	0.32	0.24	0.30
ΔMWS	0.22	0.20	0.20	0.23	0.15	0.25	0.17	0.20

These results indicate that these climate observations in the (1,0) case are much closer to those in the (1,1) cases and meanwhile the (0,1) cases are much closer to (0,0) cases. Traditionally the rain gauge is used as reference. Furthermore this study focuses on the rainfall during summer time. As a result, combining those points of view, the (0,1) can be neglected or included into (0,0) cases. Similarly, for (1,0) cases, due to its small proportion and the results from PCA, it can also be neglected or included into (1,1) cases. However, considering the rain gauge observations in the case (1,0) is qualified, so those gauge observations can also be used to estimate the marginal distributions for rain gauges.

4.4 Assumption of Independent and Identically Distributed Data

To apply the Copula-based analysis described in the last Chapter, independent and identically distributed (*i.i.d.*) data is required. Generally, in the standard approaches done by the previous work (e.g. Bardossy et al., 2006; Bardossy and Li, 2008; Serinaldi et al., 2008; Villarini et al., 2008; Bardossy and Pegram, 2009; AghaKouchak et al., 2010a and 2010b), this assumption of *i.i.d.* data was not considered. As it was shown recently by Laux et al. (2011) for daily precipitation observations and data from RCM models, this pre-requisite is not complied with and therefore the ARMA-GARCH transformation was performed to deal with this problem (Vogl et al., 2012).

Table 4.6: Ljung-Box Q-test for radar and gauge (positive pairs only) and their residuals after performing ARMA-GARCH transformation, under different lags, for selected rain gauge stations at Garmisch-Partenkirchen, Oberammergau and Wielenbach (June, July, and August of 2006 to 2007), alpha = 0.01, 1 means autocorrelated, 0 means no autocorrelation.

Station		w/o ARMA-GARCH					w/ ARMA-GARCH				
		1	5	10	15	20	1	5	10	15	20
Garmisch-P.	Radar	1	1	1	1	1	0	0	0	0	0
	Gauge	1	1	1	1	1	0	0	0	0	0
Oberammergau	Radar	1	1	1	1	1	0	0	0	0	0
	Gauge	1	1	1	1	1	0	0	0	0	0
Wielenbach	Radar	1	1	1	0	0	0	0	0	0	0
	Gauge	1	1	1	1	1	0	0	0	0	0

In this study, considering the time series for positive precipitation, the autocorrelation can still be found for both gauge and radar hourly measurements as shown in the top of Figure 4.6 and Figure 4.7 at Garmisch-Partenkirchen. As a result, the ARMA-GARCH transformation is necessary to remove the autocorrelation and heteroskedasticity parts, making the gauge/radar positive observations to be *i.i.d.*. The autocorrelation for gauge/radar residuals after performing ARMA (1,1) - GARCH (1,1) is shown in the bottom of Figure 4.6 and Figure 4.7. Furthermore, the results of Ljung-Box Q-test (Box et al., 1994), as shown in the Table 4.6 and Table 4.7 (with different

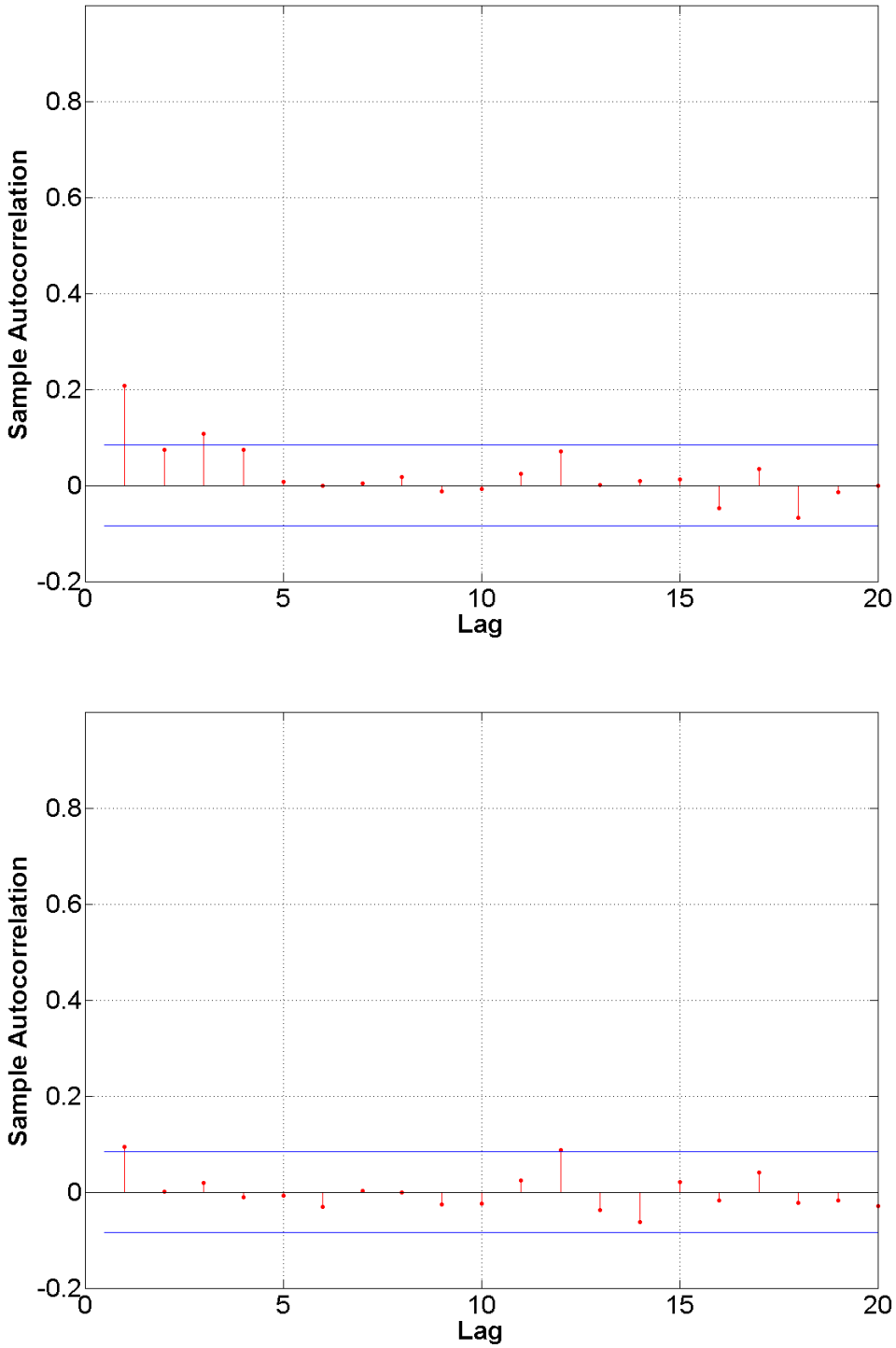

Figure 4.6: Autocorrelation for gauge observed rainfall (top) and residuals (bottom) after performing ARMA-GARCH transformation for the station at Garmisch-Partenkirchen, positive observations only, summer, 2006 to 2007.

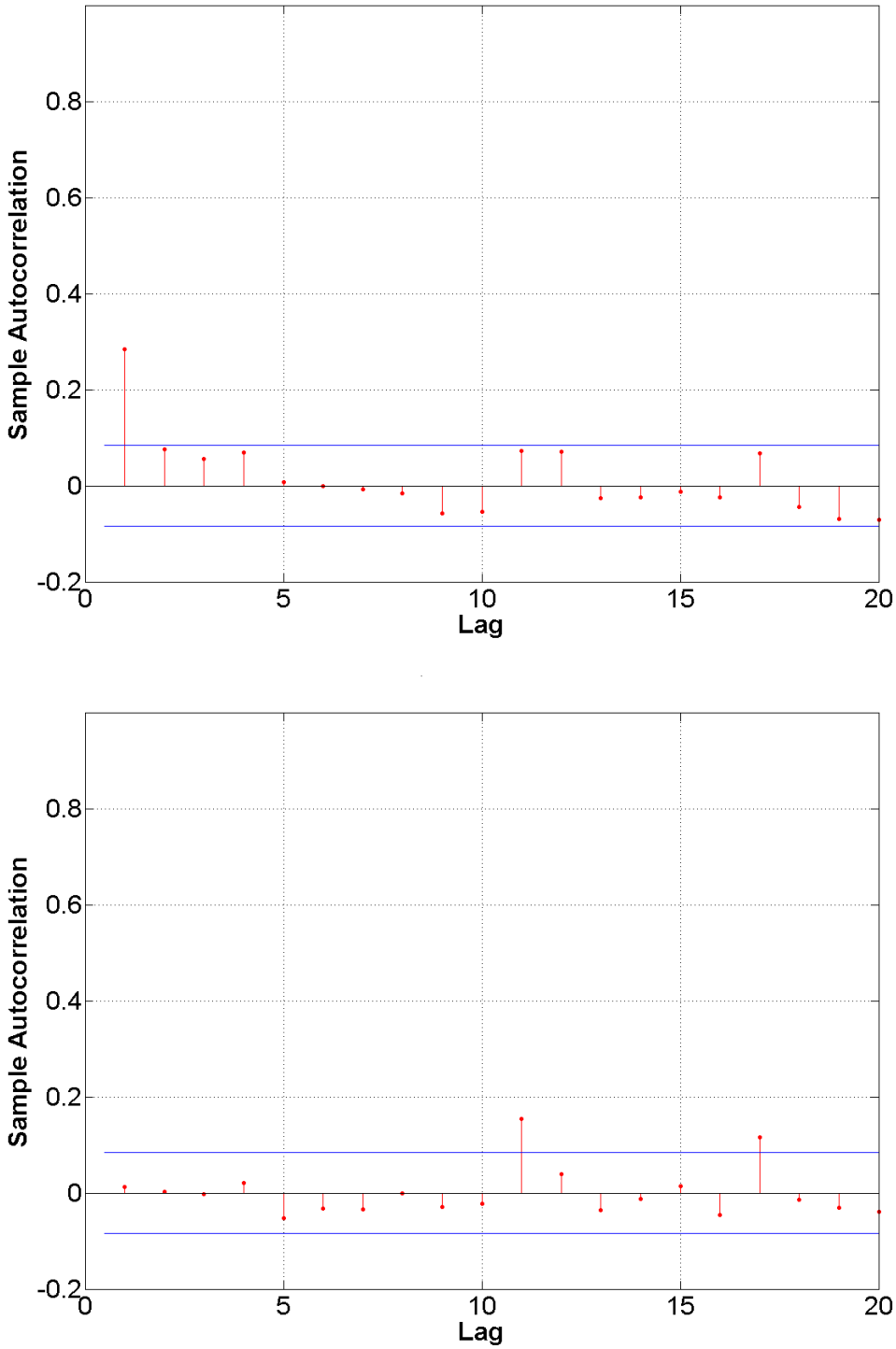

Figure 4.7: Autocorrelation for radar observed rainfall (top) and residuals (bottom) after performing ARMA-GARCH transformation at Garmisch-Partenkirchen, positive observations only, summer, 2006 to 2007.

Table 4.7: Ljung-Box Q-test for radar and gauge (positive pairs only) and their residuals after performing ARMA-GARCH, under different lags, for selected rain gauge stations at Garmisch-Partenkirchen, Oberammergau and Wielenbach (June, July, and August of 2006 to 2007), alpha = 0.05, 1 means autocorrelated, 0 means no autocorrelation.

Station		w/o ARMA-GARCH					w/ ARMA-GARCH				
		1	5	10	15	20	1	5	10	15	20
Garmisch-P.	Radar	1	1	1	1	1	0	0	0	0	0
	Gauge	1	1	1	1	1	1	0	0	0	0
Oberammergau	Radar	1	1	1	1	1	0	0	0	0	0
	Gauge	1	1	1	1	1	0	0	0	0	0
Wielenbach	Radar	1	1	1	1	0	0	0	0	0	0
	Gauge	1	1	1	1	1	0	0	0	0	0

lags and significance levels), also prove that the ARMA (1,1) - GARCH (1,1) is already sufficient to remove the autocorrelation structures existed in the original data sets. The results of Engle test (Engle et al., 1988), as shown in the Table 4.8 and Table 4.9 (with different lags and significance levels), also prove that the ARMA (1,1) - GARCH (1,1) is also sufficient to remove the heteroskedasticity structures existed in the original data sets. Therefore, in the approach w/ ARMA-GARCH transformation, the Copula parameter is estimated between the residuals to model the dependence structure between radar and gauge positive pairs (Laux et al., 2011; Vogl et al., 2012). From mathematical theoretical point of view, the *i.i.d.* data is the constraint for the application of Copulas.

However, in this way, the positive precipitation observations are selected from the original gauge or model continues time series respectively. Then all the autocorrelation analysis, Ljung-Box Q-test and Engle test are applied to these artificial series (positive observations) which is not a strict continues time series any more. This operation may have the problem that, for instance to assume that there are two exact the same rainfall events but occurred after an dry interval, e.g. 1 week. Then these two events are selected and applied with the autocorrelation analysis, Ljung-Box Q-test and Engle test. Of course, in this case, this specific selected series with positive rainfall is not *i.i.d.*. However, in fact, between them, there is a dry period for 1 week which means that these

Table 4.8: Engle test for radar and gauge (positive pairs only) and their residuals after performing ARMA-GARCH transformation, under different lags, for selected rain gauge stations at Garmisch-Partenkirchen, Oberammergau and Wielenbach (June, July, and August of 2006 to 2007), alpha = 0.01, 1 means conditional heteroscedasticity, 0 means no conditional heteroscedasticity.

Station		w/o ARMA-GARCH					w/ ARMA-GARCH				
		1	5	10	15	20	1	5	10	15	20
Garmisch-P.	Radar	1	1	1	1	1	0	0	0	0	1
	Gauge	0	0	0	0	0	0	0	0	0	0
Oberammergau	Radar	1	1	0	1	1	0	0	0	0	0
	Gauge	1	1	1	1	1	0	0	0	0	0
Wielenbach	Radar	0	1	0	1	0	0	0	0	0	0
	Gauge	1	1	1	1	1	0	0	0	0	0

two events are obviously isolated from each other and definitely not autocorrelated.

Therefore, in this thesis, both Copula-based approaches w/o and w/ ARMA-GARCH transformation are applied to investigate the sensitivity of ARMA-GARCH transformation to the final estimated precipitation fields.

4.5 Summary and Discussion

In the Alps and alpine forelands complex terrain, radar and gauge derived rainfall data from DWD were available from 2005 to 2008. Considering the rainfall estimated from a weather radar, the source of error are manifold. Generally, they can be classified into 4 different classes as described below.

1. Errors due to the radar system if the radar system losses and the antenna gain are not known precisely.

2. Radar measurements not related to rainfall, such as ground clutter.

3. Non-representative sampling space-the most important source of error is the fact that usually the radar pulse volume is located far above the area of interest (Seo et al., 1999)

Table 4.9: Engle test for radar and gauge (positive pairs only) and their residuals after performing ARMA-GARCH transformation, under different lags, for selected rain gauge stations at Garmisch-Partenkirchen, Oberammergau and Wielenbach (June, July, and August of 2006 to 2007), alpha = 0.05, 1 means conditional heteroscedasticity, 0 means no conditional heteroscedasticity.

Station		w/o ARMA-GARCH					w/ ARMA-GARCH				
		1	5	10	15	20	1	5	10	15	20
Garmisch-P.	Radar	1	1	1	1	1	0	0	0	0	1
	Gauge	0	0	0	0	0	0	0	0	0	0
Oberammergau	Radar	1	1	1	1	1	0	0	0	0	0
	Gauge	1	1	1	1	1	0	0	0	0	0
Wielenbach	Radar	1	1	1	1	1	0	0	0	0	0
	Gauge	1	1	1	1	1	0	0	0	0	0

4. Indirect measurement of rainfall, e.g. the Z/R relationship.

The relative importance of the error sources is often difficult to quantify and may differ under various conditions and amount to systematic and random differences from rain gauge measurements as large as 100 percent or more (Smith et al., 1996). On the other hand, even though the rain gauge can provide ground accurate direct rainfall measurements. However, it is difficult to extend to spatial rainfall field by using the point gauge observations.

The radar rainfall maps can provide spatial information for the rainfall and meanwhile the accurate point value can be observed from the rain gauges at grid scale. A new method is somehow urgent to combine radar and gauge data in a reasonable way, or even with the ability to further combine the new rainfall measurement device named microwave attenuation. This will be described in the next chapters.

In this chapter, at the third section, the 4 different situations in gauge and radar pairs are analysed, checked and discussed. Some statical results of those (0,0), (0,1), (1,0) and (1,1) cases are compared and discussed for different seasons from 2005 to 2008. Additionally, the PCA and factor analysis are performed on the temperature, relative humidity and mean wind speed to further reveal the more detailed information in those 4 cases. The final conclusion can be made that the rain gauge can be used as

the reference to determine rainy or not for each time step. This means that the (0,1) case can be included into (0,0) and (1,0) case can be included into (1,1) case, especially at summer. In this case, the gauge/radar pairs' structure is also simplified so that the main effort can be put on positive pairs at summer time.

In the fourth part, the *i.i.d.* condition is discussed. Although, the *i.i.d.* assumption was neglected in the previous research applications, according to the results from the autocorrelation function, Ljung-Box Q-test and Engle test, the ARMA-GARCH transformation is still necessary to apply for the hourly precipitation observed by gauge and radar in this study area. As a result, in this thesis, the simulation results from both cases w/o and w/ ARMA-GARCH transformation are presented and compared in the following chapters.

5

New Precipitation Data - Microwave Attenuation

5.1 Introduction

In contrast to the point scale rain gauge and biased spatial radar, as well as their costly installation and operation, the need for additional rainfall measurements inspired the use of microwave link attenuation, which has been proven to provide accurate rainfall estimation at the near-surface level. The physical principle of this new device is due to the fact that rain drops considerably absorb and scatter electromagnetic radiation in the microwave region (wavelength between 0.3 cm and 30 cm, corresponding to frequencies between 1 GHz and 100 GHZ) (e.g. Chwala et al., 2012). In the past, several experiments with purpose-built MW-links, mostly using two frequencies, were conducted to measure line integrated precipitation (e.g. Holt et al., 2003; Rahimi et al., 2004; Minda et al., 2005; Leijnse et al., 2007a) and its drop size distribution (DSD) (e.g. Rincon and Lang, 2002). However, in those experiments, the needed equipment was still too expensive for a wide areal coverage (e.g. Chwala et al., 2012). To overcome this, Messer et al. (2006) showed that it is possible to use attenuation data from existing commercial MW-link networks operated by cell phone providers. Those networks exist wherever a cell phone network exists. Thus it is available in many regions, even in areas with a coarse station network like in mountainous regions or developing countries. Using the whole network as data source offers the opportunity to generate countrywide interpolated rainfall maps (e.g. Zinevich et al., 2008).

In this chapter, after the introduction part, the physical background of the MW-link is introduced in detail. Then both the theoretical and empirical derivations from microwave attenuation to rain rate are presented in detail. Afterwards, some results are shown derived from data acquired at 2/5 MW-links in the Alps and alpine forelands in southern Germany for the period July 2010 to October 2010. This includes a new wet/dry estimation algorithm which uses spectral time series analysis to improve the baseline determination for attenuation data, followed by the resulting rain rates and then compared to the corresponding gauge and radar derived quantities. Finally, the conclusion of this chapter is given.

5.2 Physical Background

5.2.1 Drop Size Distribution

The drop size distribution (DSD, quantity symbol N) represents the probability density of equivolumetric drop diameter D being in the unity volume. The product $N(D)\mathrm{d}D$ gives the number of rain of the diameter between D and $D + \mathrm{d}D$ in the unity volume. Only rain drops of diameters below 7 mm can exist for physical reasons (Fiser et al., 2010).

The Gamma and Exponential distribution are the most frequently used analytical approximations of the DSD because of their satisfactory correspondence with the typical drop size distribution shape in the majority of experimental samples (Fiser et al., 2010). Many other DSD models can be found in the literature, for instance the log-normal model (e.g. Ajayi and Kozu, 1999). Note that other factors like the rain type, time of integration and others influence the analytical DSD modelling (Fiser et al., 2010). The Gamma model of DSD is listed below

$$N(D) = N_0(D)D^{\mu}exp(-\Lambda D) \tag{5.1}$$

where, D [mm] is the rain drop diameter; $N(D)$ $[m^{-3}mm^{-1-\mu}]$ is the number of drops per unit volume per drop diameter interval ($\mathrm{d}D$); N_0 $[m^{-3}mm^{-1-\mu}]$ is the intercept parameter of DSD; Λ $[mm^{-1}]$ is the slope parameter; μ [-] is shape of the DSD.

The relatively simpler Exponential distribution is also used as shown in the following expression. When the parameter μ equals to zero, the Exponential DSD model becomes

the Gamma DSD.

$$N(D) = N_0(D)exp(-\Lambda D) \tag{5.2}$$

The parameter λ in this model is directly related to the rain rate R:

$$\Lambda \approx \alpha R^{\beta} \tag{5.3}$$

5.2.2 Mie Scattering

The Mie scattering theory was developed by Mie et al. (1908). The scattering function f (f_f for forward and f_b for backward scattering) for spherical dielectric particles is given by the formulas below (Fiser et al., 2010).

$$f_f = \frac{-j\lambda^3}{\pi^3 D^2}[\sum_{i=1}^{\infty}(2n+1)(a_n + b_n)\,]^* \tag{5.4}$$

$$f_b = \frac{-j\lambda^3}{\pi^3 D^2}[\sum_{i=1}^{\infty}(-1)^{n+1}(2n+1)(a_n - b_n)\,]^* \tag{5.5}$$

where, λ denotes the vacuum wavelength of the electro-magnetic radiation; j is the imaginary unit and D the diameter of the spherical drops; $*$ is symbol for conjugate imaginary numbers. The coefficients a_n and b_n depend on the complex relative refractivity $\varepsilon_r = \varepsilon/\varepsilon_0$ of the material (here, rain water) and on the diameter D of the scattering sphere. a_n and b_n are the Mies coefficients.

In the case if the rain drop shape can be accepted as being spherical (for larger rain drops it is not true), Mie scattering can be used to study the frequency and temperature properties of rain attenuation. Mie scattering does not enable to compute depolarisation and angular dependencies, but with no frequency limitation like in the Rayleigh scattering computation case (Fiser et al., 2010).

5.3 Attenuation and Rain Rate

This section is a revised version of the publications by Olsen and Rogers (1978) and Rogers and Olsen (1976).

5.3.1 Theory

Assumed to be a plane wave, or furthermore an isotropic or anisotropic spherical wave, the general relation for the average field transmitted through a random discrete-scatterer medium of length l and bulk refractive index $\eta \cong 1$ can be formulated as below (e.g. Olsen and Rogers, 1978; Rogers and Olsen, 1976; van de Hulst et al., 1957; Ishimaru and Lin, 1973).

$$\langle E(\overline{r},t)\rangle = exp(-jkl(\eta-1))E_i(\overline{r},t) \tag{5.6}$$

$$E_i(\overline{r},t) = E_i(\overline{r})exp(-j2\pi ft) \tag{5.7}$$

$E_i(\overline{r},t)$ is the field of wavelength λ and direction $\overline{k}(k = 2\pi/\lambda)$ incident on the medium. As a result, the specific attenuation A (both absorption and scattering) in db/unit distance can be written as

$$A = 20kIm(\eta)/ln(10) \tag{5.8}$$

And the bulk index for a medium of sparsely distributed scatterers such as rain is given by

$$\eta = 1 - j(2\pi/k^3)\int_0^\infty S(0,D)N(D)\,\mathrm{d}D \tag{5.9}$$

where $S(0,D)$ is the forward scattering amplitude and $N(D)\mathrm{d}D$ is the number density of raindrops with equivalent diameter D in the interval $\mathrm{d}D$, the same as introduced in last section about DSD. By substituting Eq. (5.9) in Eq. (5.8), the most frequently used expression for A in terms of the extinction cross section $Q_t(D) = (4\pi/k^2)Re[S(0,D)]$ is obtained:

$$A = 4.343\int_0^\infty Q_t(D)N(D)\,\mathrm{d}D \tag{5.10}$$

So, the attenuation A is dependent on the two different parts. One is the scattering part $N(D)\mathrm{d}D$, which is the drop size distribution. Another part is the Mie scattering, referred to $S(0,D)$, as the spherical particle can be assumed and the microwave length λ is large enough comparing to the rainfall drop size D.

For the Mie scattering part, the infinite series expression for the forward scattering amplitude of a spherical particle can be approximated by the first few terms (van de Hulst et al., 1957).

$$S(0,D) = 1/2[3(a_1+b_1) + 5(a_2+b_2) + 7(a_3+\cdots) + \cdots] \tag{5.11}$$

where a_n, b_n,etc,., are the Mie coefficients. Using the expressions for the first few Mie coefficients to order x^8 in the size parameter $x = \pi D/\lambda$, it follows

$$S(0, D) = jx^3(M_1 + M_2x^2 + M_3x^3 + M_4x^4 + M_5x^5 + \cdots) \tag{5.12}$$

where $M_1 = (m^2 - 1)/(m^2 + 2)$,..., and m the frequency and temperature dependent refractive index of water. The details are referred to Penndorf et al. (1962). The truncated series Eq. (5.12) for $S(0, D)$ can be substituted into the integrand of Eq. (5.9) and Eq. (5.10), along with a suitable theoretical form for the drop size distribution $N(D)$, as mentioned in previous part about DSD, which can be either Exponential or Gamma form. In case of choosing Exponential distribution for DSD, the following form can be generated

$$A = a^{'}R^{b^{'}}(1 + c_2f^2R^{2\beta} + c_3f^3R^{3\beta} + c_4f^4R^{4\beta} + c_5f^5R^{5\beta} + \cdots) \tag{5.13}$$

where

$$a^{'} = 60\pi N_0 fIm(M_1)/ln(10)c\alpha^4, b^{'} = 4\beta \tag{5.14}$$

and

$$c_n = (n + 1)!/3!(\pi/c\alpha)^n Im(M_n)/Im(M_1), n = 2, 3, 4... \tag{5.15}$$

with c the speed of light. If the widely accepted empirical form $\Lambda = \alpha R^\beta$ as used in the exponential distribution, also in Gamma case, then the general form can be obtained as

$$A = a^{'}R^{b^{'}}(1 + c_2f^2R^{2\beta/q} + c_3f^3R^{3\beta/q} + c_4f^4R^{4\beta/q} + c_5f^5R^{5\beta/q} + \cdots) \tag{5.16}$$

where now

$$a^{'} = 10\pi^2\Lambda_1 f\Gamma(p + 4/q)Im(M_1)/ln(10)qc\alpha^{p+4/q}, b^{'} = (p + 4)\beta/q \tag{5.17}$$

and

$$c_n = \Gamma(p + n + 4/q)\Gamma^{-1}(p + 4/q)(\pi/c\alpha^{1/q})^n Im(M_n)/Im(M_1), n = 2, 3, 4... \tag{5.18}$$

Furthermore, if all Mie coefficients a_n and b_n had been retained in the series for $S(0, D)$ and if all powers of x had been retained in the expansions for these coefficients, then an infinite series of the following form can be obtained

$$A = a^{'}R^{b^{'}}(1 + \sum_{n=2,\infty} c_nf^nR^{n\beta/q}) \tag{5.19}$$

5.3.2 Empirical

Since M_n and therefore c_n are unavailable beyond $n = 5$, the region of convergence of (5.19) cannot be determined. The empirical relationship shown in Eq. (5.20) between A in [dB/km] and R in [mm/hour] remains the better approximation to the full Mie solution (e.g. Olsen and Rogers, 1978).

$$A = aR^b \tag{5.20}$$

where both a and b are assumed to be independent of R and dependent on the specific frequency, drop size distribution and temperature (e.g. Olsen and Rogers, 1978). For different rainfall climatologies and DSD there are several sets of the values a and b available, both empirically and theoretically derived ones. What they all have in common is, that b is close or equal to one in the frequency range from 25 GHz to 40 GHz. Hence, the power law in Eq. (5.20) is almost linear, which is an important fact, since any non linearity leads to errors of the line integrated precipitation because of the inhomogeneous distribution of rain fall along the MW-link path (Leijnse et al., 2008). In the following parts, the a and b values from the recommendation by ITU (2003) and do an shape preserving spline interpolation for the values in between the ones listed.

5.4 MW-Link Derived Precipitation

5.4.1 MW-Link Distribution

In this study region the access to several MW-links are available from Ericsson GmbH which are parts of a German cell phone network. According to data availability problem of the C-band experimental weather radar in Hohenpeissenberg, only the two MW-links in Figure 5.1 can be covered by the radar field, which are the MW-link 1 from Hohenpeissenberg (hop2) to Weilheim (wh0) and MW-link 2 from Hohenpeissenberg (hop2) to Murnau (murn1). The geographic and technique description of these two MW-links are listed in Table 5.1.

5.4.2 Wet/Dry Determination

It is the first step to distinguish the wet and dry periods to generate precipitation from microwave link attenuation data. The spectra analysis based algorithm for wet/dry de-

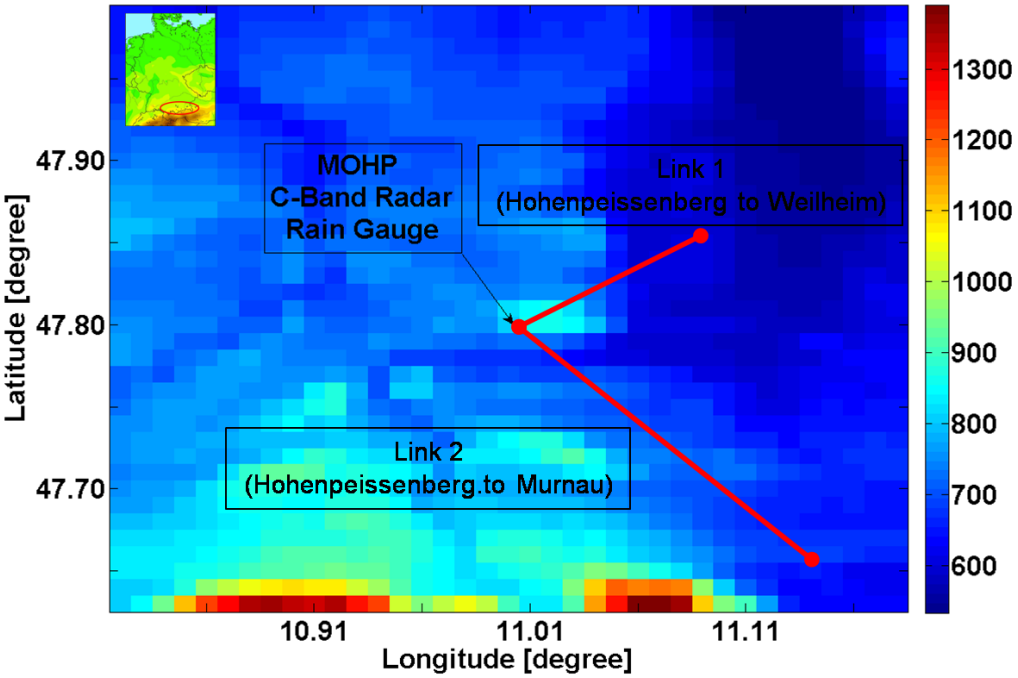

Figure 5.1: Overview of the study region in southern Germany showing data sources - radar and MW-links. The unit of the color bar is in [m.a.s.l].

termination of MW-Links is taken from Chwala et al. (2012) and the brief introduction is given in the following.

For each time step t a short section of the link attenuation data is generated as following

$$L(t) = \{L_k | k \in (t-l, \cdots, t+l)\} \tag{5.21}$$

with length of $2l$, windowed by a Hamming window w from which we calculate the Fourier transform

$$F(f,t) = FFT(wR(t)) \tag{5.22}$$

via FFT. As we are only interested in the spectrum of the amplitudes we use the power spectral density

$$P(f,t) = \frac{|F(f,t)|^2}{F_s \sum_0^l w} \tag{5.23}$$

As it is hard to classify the deviations from the mean spectrum of the dry period, for each point in time t, the a normalized spectrum (calculated by division through a

Table 5.1: The geographical and technical description of the 2 MW-links

ID /Unite	MW-Link	Length [km]	Altitude [m.a.s.l]	Frequency [GHz]	Polarization
1	hop2-wh0	10.2	420	18.7	Vertical
2	hop2-murn1	17.4	300	15.0	Vertical

mean dry spectrum $P_{meandry}$) is involved as listed below:

$$P_{norm}(f,t) = \frac{P(f,t)}{P_{meandry}(f,t)} \tag{5.24}$$

Figure 5.2 shows the three spectra for the three different atmospheric conditions in both original and normalized forms, revealing that, it is necessary to take the difference between $P_{sumlow}(t)$ and $P_{sumhigh}(t)$.

$$P_{sumdiffer}(t) = P_{sumlow}(t) - P_{sumhigh}(t) \tag{5.25}$$

Here, $P_{sumlow}(t)$ and $P_{sumhigh}(t)$ are defined as following:

$$P_{sumlow}(t) = \sum_{f=f_{low1}}^{f_{low2}} \frac{P_{norm}(f,t)}{N_{low}} \tag{5.26}$$

$$P_{sumhigh}(t) = \sum_{f=f_{high1}}^{f_{high2}} \frac{P_{norm}(f,t)}{N_{high}} \tag{5.27}$$

with $f_{low1} < f_{low2} < f_{high1} < f_{high2}$ and N_{low}, N_{high} being the number of frequencies used in each sum. Once their difference exceeds a certain threshold σ, this period can be identified as wet.

$$t = \begin{cases} wet & \text{if } P_{sumdiffer}(t) > \sigma \\ dry & \text{if } P_{sumdiffer}(t) \leq \sigma \end{cases}$$

Within the wet periods the baseline L_{base} of the RSL is assumed to keep constant at the level of the last value $t_{lastdry}$ of the last dry period. So, the baseline is defined as following.

$$L_{base} = \begin{cases} L_{t_{lastdry}} & \text{if wet} \\ L_t & \text{if dry} \end{cases}$$

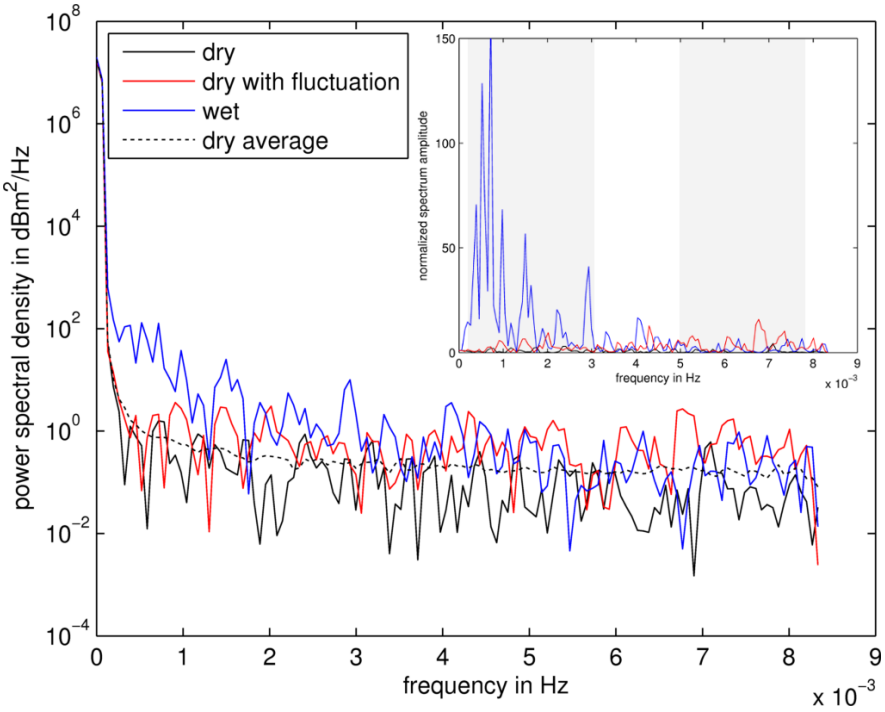

Figure 5.2: Typical spectra for 256 minute snippet (with a Hamming window) from a RSL time series for different atmospheric conditions. The inset shows the spectra normalized with the average dry spectrum by division. Shaded areas mark the frequency ranges (low and high) which are used to compare the amplitude sums to decide whether the snippet is from a wet or a dry period. Deviation from the mean dry spectrum is largest for the low frequency part of the wet spectrum. Note that for the dry with fluctuation-spectrum the observable deviation is highest in the high frequency part (Chwala et al., 2012).

As a result, the attenuation is calculated as

$$A(t) = L_{base}(t) - L_t \tag{5.28}$$

which is always zero during the dry periods. If the attenuation and thus the rain rate get negative, which sometimes happens at the end of a rain event when the baseline is set to low, the rain rate is set to zero.

5.4.3 MW-Link vs Gauge and Radar

By applying the wet/dry estimation mentioned in the previous section, the MW-link derived precipitation can be calculated and compared to the precipitation from gauge and radar as shown in the top of Figure 5.3. Note that the radar value is the weighted

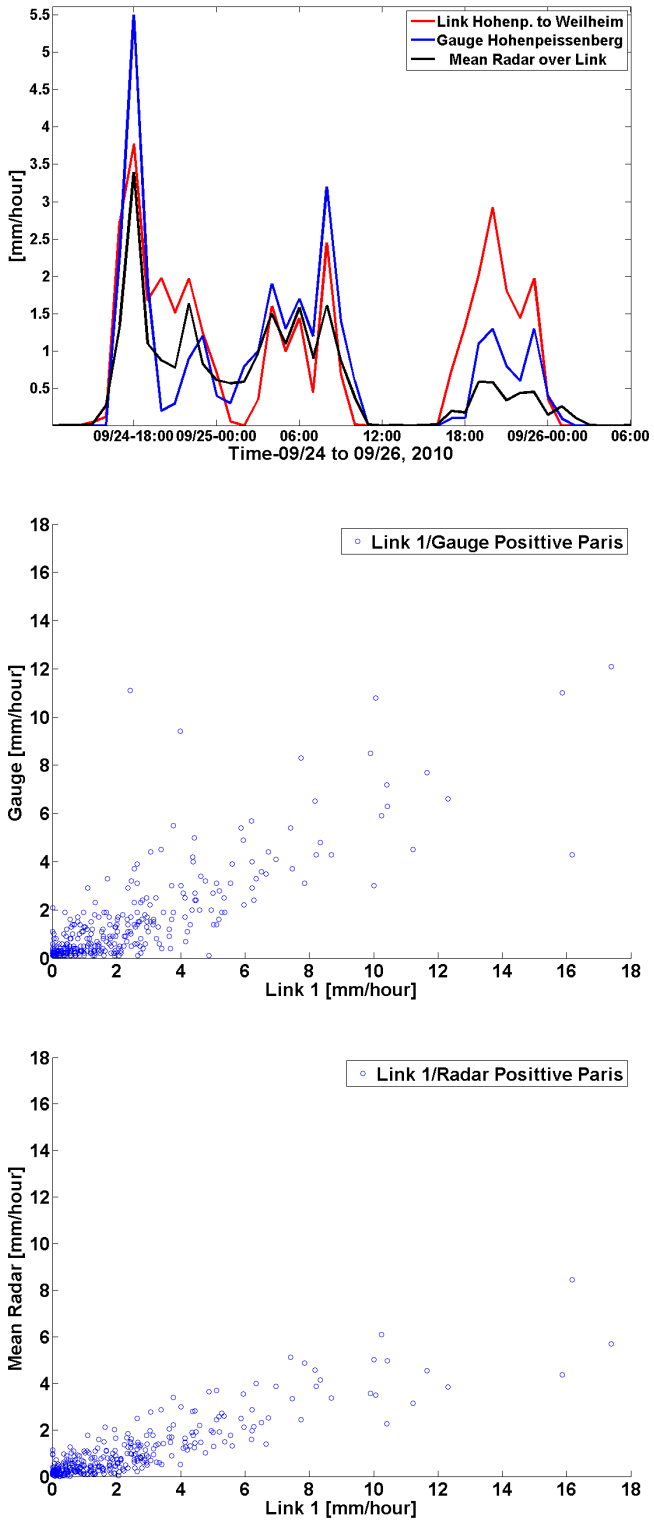

Figure 5.3: MW-link derived rainfall versus radar and gauge at Hohenpeissenberg, the time series plot for exemplary events and two scatter plots (from top to bottom), June to October, 2010.

mean value along the grids beneath the MW-link from Hohenpeissenberg to Weilheim and the rain gauge is also located at the top of mount Hohenpeissenberg but western to the MW-link. The resulting MW-link observed rainfall illustrates a highly correlated structure comparing to the corresponding gauge and radar measurements. The correlation coefficients between the MW-link and gauge/radar can be 0.80 and 0.87 and the RMSE value are 1.91 and 2.27. This also proves that generally the wet/dry determination method works well. As a result, it can be concluded that the precipitation derived from microwave link attenuation is qualified and can be a good complement for gauge and radar.

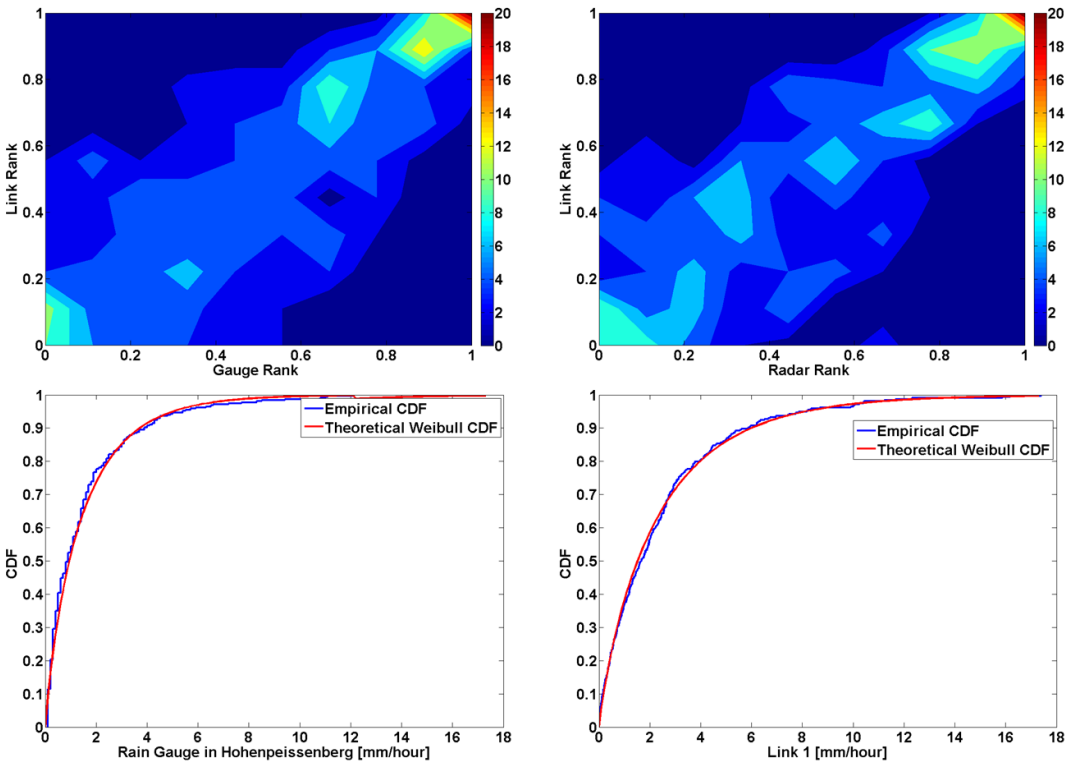

Figure 5.4: Density of the empirical Copula derived from MW-link 1/radar and MW-link 1/gauge positive pairs as well as their marginal distributions (top to bottom), Hohenpeissenberg, June to October, 2010.

Some problems still exist, such as errors from wet antenna, simplified Mie scattering and others. For instance, as also can be seen from the middle and bottom of Figure 5.3, the MW-link can under-/over- estimate the rainfall compared to the gauge due to the a and b coefficients even though the correlation coefficients can be high. In the

ideal case, a and b should be different under variant DSD and temperature, requiring a further development of better a and b values. An experiment to gather DSD spectra in the vicinity of the used MW-links to calculate improved a and b values for this study region is ongoing at the moment (Chwala et al., 2010). However, the proposed Copula based approach can overcome this a and b quantization problem by using the rank value instead of the absolute value as shown in Figure 5.4. The empirical Copulas at the top show that the MW-link/gauge and MW-link/radar are both highly correlated and the upper/lower tail dependence structures can also be found for both. Furthermore, the marginals also fit well as shown in the bottom for gauge and MW-link. The detailed information about the Copula can be found in the next Chapter.

5.5 Summary and Discussion

In this Chapter, the physical background, the A/R relationship and some preliminary results are presented and to prove that the MW-link is a good complement for radar and gauge. According to the comparisons between the MW-link and gauge/radar, the MW-link derived precipitation is highly correlated to both gauge and radar measurements. However, due to the problems such as dynamic baseline and a/b values, the absolute precipitation value calculated from the microwave attenuation can over/underestimated the rainfall. The dynamic baseline and a/b values determination methods are required.

The proposed Copula-based analysis can just focus on the rank transformed values and can be the suitable method to further include the MW-link observed precipitation together with gauge and radar. In Chapter 10, the approaches, to further assimilate precipitation information from the MW-links, are developed based on the radar and gauge data integration methods. The resulting precipitation fields can be improved with the integration of MW-link derived rainfall.

6

Point Wise Data Integration of Gauge and Radar

In this chapter, the point wise Copula-based statistical approach is introduced to combine radar and gauge observations. Information from rain gauge and the corresponding radar within the same grid cell is assimilated through studying the dependence structure between them. Then the pseudo observations are simulated conditioned on the given radar observations. The structure of this chapter is as following. Firstly, in section 6.1, a briefly introduction is given. Afterwards, marginal distributions for both gauge and radar are estimated in section 6.2. Then, in section 6.3 and 6.4, the Copula-based analysis for radar/gauge positive pairs are presented, including both of the results w/o and w/ ARMA-GARCH transformation, as well as their validations separately. Finally, the discussion and conclusion is given in section 6.5.

6.1 Introduction

Generally, rain gauge observations are used as the 'ground truth', while the radar can provide spatial precipitation pattern regarded as being superior to an interpolated rain gauge field. However, in reality, the situation is much more sophisticated. As already described in Chapter 4, the discrepancies between radar and rain gauge are mainly due to the different spatio-temporal sampling properties. Therefore, the question is now how to compare and combine the gauge and radar observations within the same grid cell in a reasonable way. One approach was done by using rain gauge data to

optimize the radar Z/R relationship. The DWD (DWD, 2001) introduced a rainfall type dependent Z/R relationship trying to mitigate radar rainfall under estimation at far ranges. Marx et al. (2007) developed 3 classes of Z/R relationships are developed based on the radar reflectivity. Apart from those, Matsoukas et al. (1999) developed an alternative radar gauge fusion method based on Artificial Neural Network. Kyriakidis et al. (2001) proposed a new method to improve radar data with rain gauge observations by using Kalman filtering and variational analysis in a self-adaptive way to estimate the system noise variance and the observation noise variance. As already introduced in the Chapter 1, the tacit assumption of Gaussian behaviour is restricting the performance of standard approaches, as many studies showed that the interdependence in hydrological or meteorological data sets is usually more complex (e.g. Gomez-Hernandez and Wen, 1998; Bárdossy et al., 2006; Bárdossy and Li, 2008). Recently, the Copula-based models are widely used in the estimation of rainfall (e.g. Zhang and Singh, 2008; Serinaldi et al., 2008; Villarini et al., 2008), carried out in the bivariate framework to describe dependency between two variates.

In this chapter, the main effort is focused on the gauge/radar positive pairs and the reason for this has already been demonstrated in Chapter 3. The Copula-based simulation results from both approaches, w/o and w/ ARMA-GARCH transformation, are presented and compared.

6.2 Marginal Distributions

As explained in Chapter 3, for the probability integral transformation, the univariate marginal distributions are required both for gauge and radar. Therefore, following the first step of the Copula-based analysis, the appropriate theoretical marginal distribution functions have to be selected and estimated for all radar grid cells and rain gauges. This step keeps the same for both of the cases w/o and w/ ARMA-GARCH transformation.

For both radar and gauge positive observations, the Normal, Exponential, Gamma and the Weibull distribution, as already detailed described in Chapter 3, were considered in this study. Then, the AIC and BIC values were calculated for each radar grid and each gauge station to decide for the best fit. The results for the rain gauges located at Garmisch-Partenkirchen, Oberammergau and Wielenbach and their corresponding radar grids (summer, 2006 to 2007) are listed in Table 6.1. The Weibull distribution

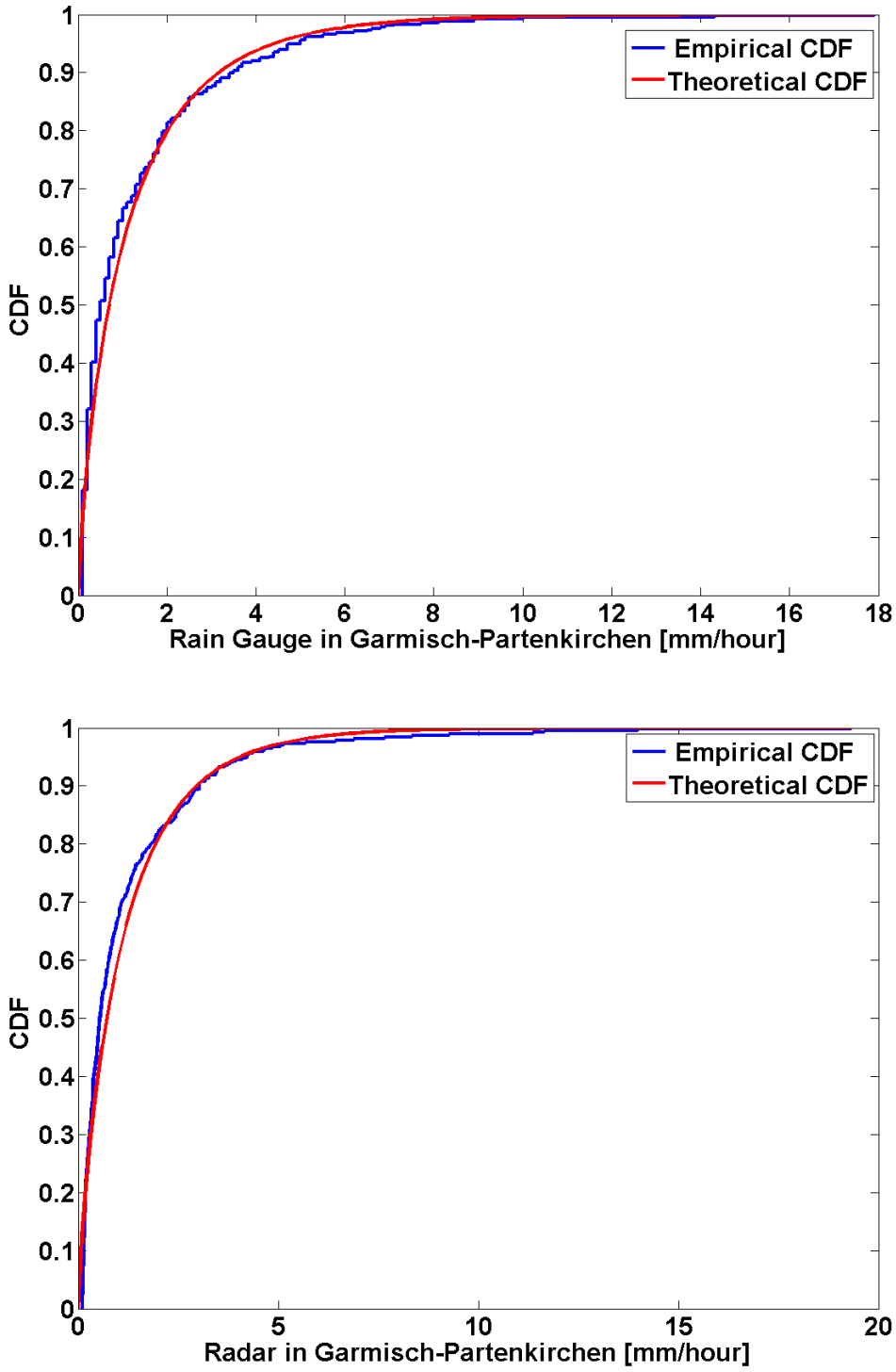

Figure 6.1: CDF of empirical and estimated theoretical marginal distribution (Weibull distribution) for the rain gauge at Garmisch-Partenkirchen (top) and the corresponding radar grid (bottom), positive observations, summer, 2006 to 2007.

Table 6.1: AIC and BIC calculated for rain gauges and the corresponding radar measurements (positive observations, ≥ 0.1 [mm/hour]) for selected locations at Garmisch-Partenkirchen, Oberammergau and Wielenbach, for different univariate theoretical distribution functions, summer, 2006 to 2007.

Station/Distribution		Normal	Exponential	Gamma	Weibull
Garmisch-Partenkirchen	AIC (Radar)	3137	1809	1783	**1740**
	BIC (Radar)	3149	1815	1794	**1751**
	AIC (Gauge)	2215	1294	1274	**1251**
	BIC (Gauge)	3139	1895	1871	**1842**
Oberammergau	AIC (Radar)	3153	1962	1918	**1880**
	BIC (Radar)	3164	1968	1929	**1891**
	AIC (Gauge)	2958	1835	1811	**1787**
	BIC (Gauge)	3013	1861	1835	**1807**
Wielenbach	AIC (Radar)	2222	1420	1417	**1405**
	BIC (Radar)	2233	1425	1427	**1416**
	AIC (Gauge)	2060	1171	1138	**1110**
	BIC (Gauge)	2590	1497	1467	**1434**

always shows the best fitting results. The GoF test results are similar for all the other gauge stations (results are not listed in detail here) and radar observations (more than 95 percent of all radar grid cells suggest for the Weibull distribution). However, only very small differences (only around 3 percent or even less) can be found among the Weibull, Gama and Exponential distributions. Therefore, the Exponential distribution with one parameter can also be another choice and will be used as introduced in Chapter 9. Note that all the marginal distributions are only fitted to the positive observations (≥ 0.1 [mm/hour]) in the corresponding time series.

A comparison of the empirical and the fitted Weibull distribution's CDF for both radar and gauge at Garmisch-Partenkirchen is shown in Figure 6.1. For both of them, it is clear that the estimated Weibull distributions match with the corresponding empirical CDF well. Even though small under-/over-estimations can still be found at regions with small/large values (comparing to 5 mm/hour) for radar. It can be concluded that the Weibull can model both gauge and radar marginal best among those 4 theoretical

distributions. Therefore, throughout this study (except for special note), the Weibull distribution is used to represent for the marginal distributions for gauge and radar.

In different seasons, the different temperatures, pressures, weather types and etc., can lead to the different precipitation types, i.e. rain to snow - liquid to solid. Therefore, marginal distributions for both gauge and radar can also be with variations. As supported by the findings in the Figure 6.2 and Figure 6.3, for radar and gauge respectively, the estimated Weibull's parameters are different in different seasons, showing the clear season dependent pattern for each locations.

In general, both for radar and gauge, the behaviour of Weibull's scale parameter is clear, with the highest value in summer and lowest value in winter. This implies the scale parameter is influenced by the temperature also with the highest in summer and lowest in winter. However, for Weibull's shape parameter of radar, the pattern is not so straightforward. Especially at the location of Garmisch-Partenkirchen, the highest one is found at autumn and the second highest in winter. On the other hand, considering the Weibull's shape parameter for gauge, it is clear again with the highest in winter and lowest in summer, just opposite of the Weibull's scale parameter. So, the conclusion can be made that the marginal distributions for radar and gauge are influenced by the seasonal/temperature changing. Further discussion about this temperature dependent behaviour for marginal distributions can be found in the Chapter 9.

6.3 Copula Analysis without ARMA-GARCH Transformation

In this section, the procedure of searching for the best fit Copula $C(F_R(r), F_G(g))$ to model the dependence structure between radar/gauge positive pairs is introduced at first, following the standard ways w/o ARMA-GARCH transformation to neglect the *i.i.d.* assumption. Based on the estimated marginal distribution and Copula dependence function, the Copula-based pseudo observations are simulated conditioned on the given radar measurements within one radar grid cell.

6.3.1 Copula Fitting

After using the univariate marginal distributions to transform the original data to the rank space, the empirical Copulas are calculated and then the dependence structure can

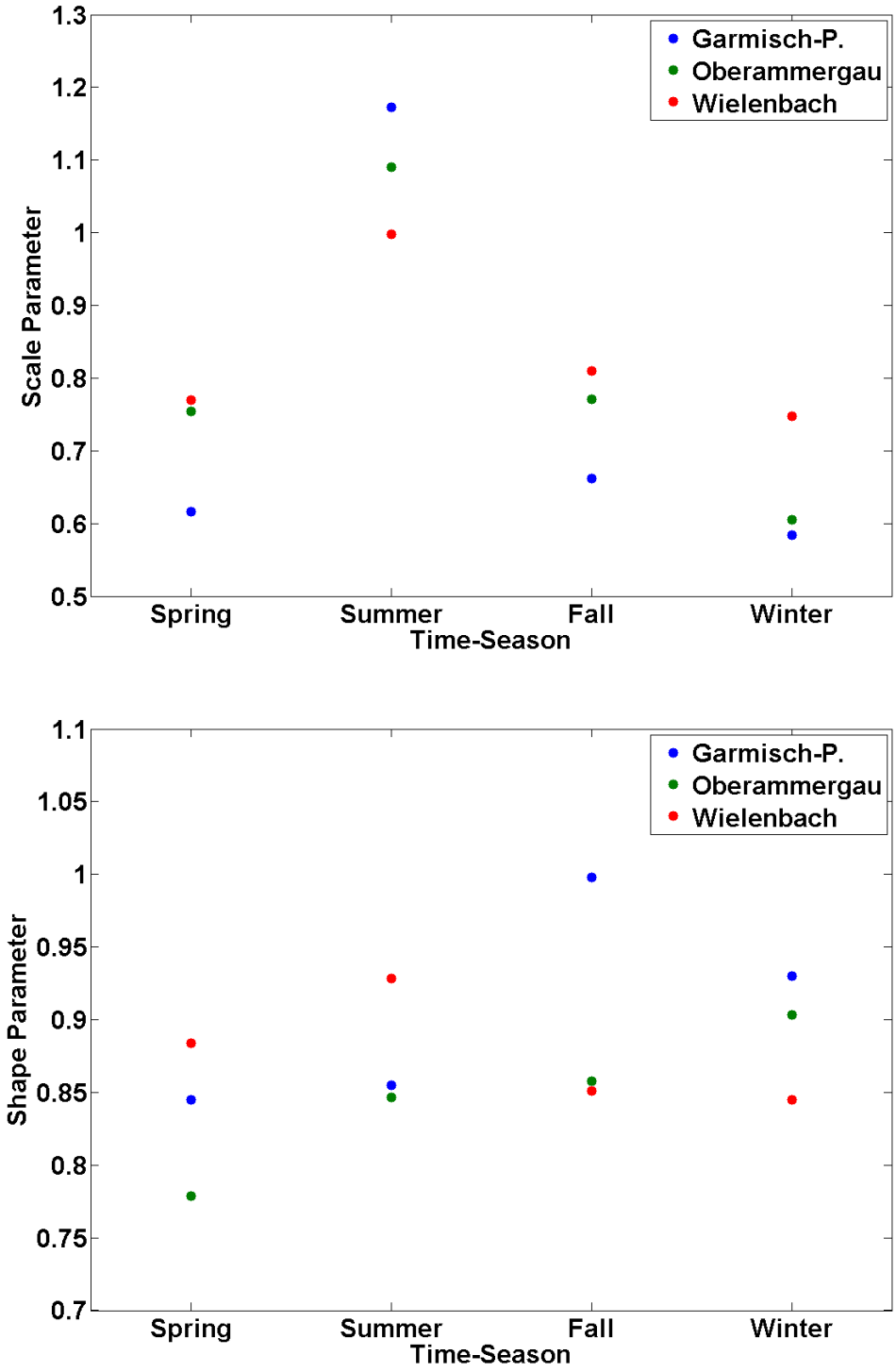

Figure 6.2: Estimated parameters of the Weibull distributions (scale/shape, top/bottom) for radar data (positive observations) at Garmisch-Partenkirchen, Oberammergau and Wielenbach, in different seasons, 2005 to 2008.

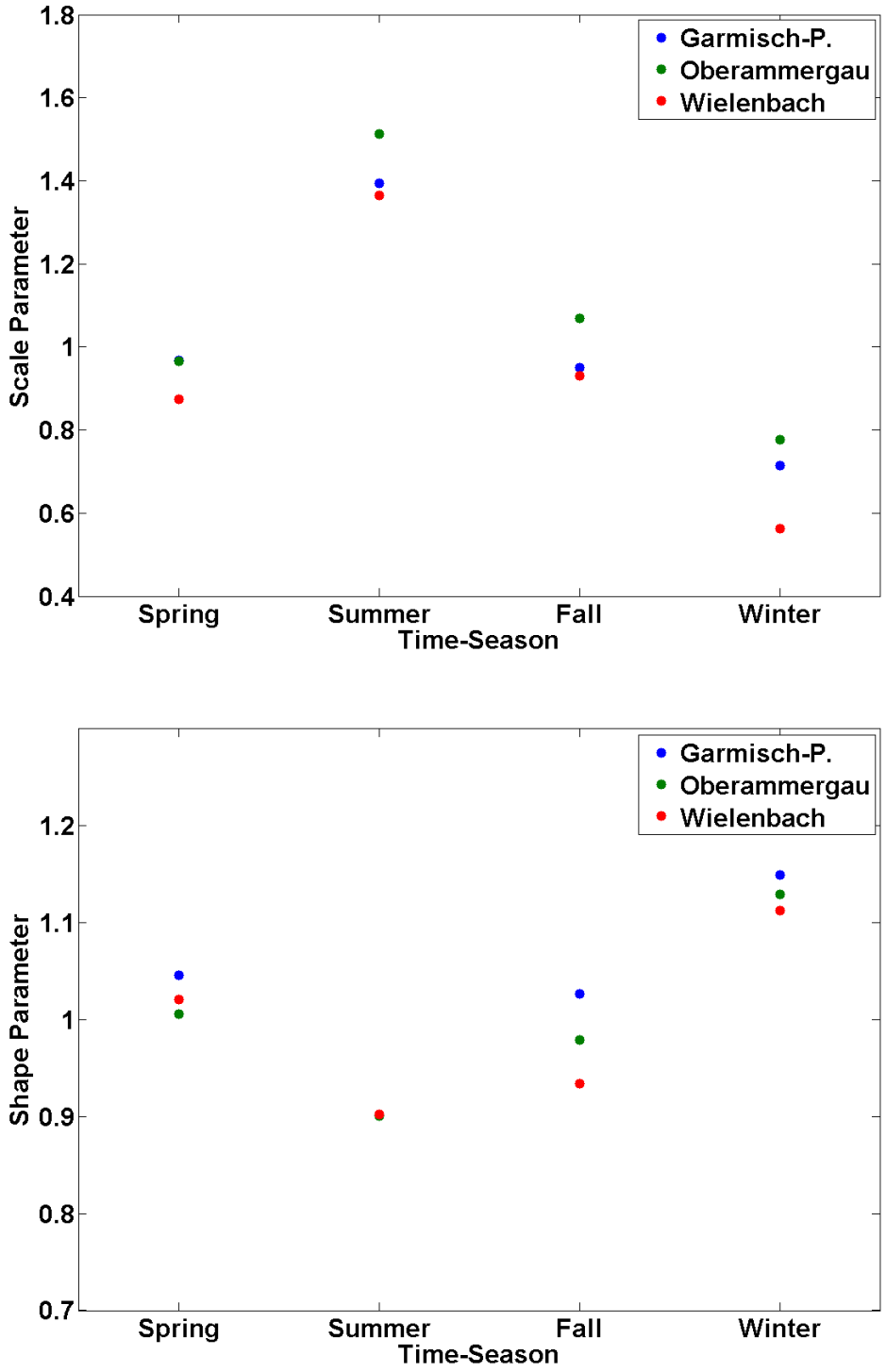

Figure 6.3: Estimated parameters of the Weibull distributions (scale/shape, top/bottom) for gauge data (positive observations) at Garmisch-Partenkirchen, Oberammergau and Wielenbach, in different seasons, 2005 to 2008.

be investigated between the radar/gauge positive pairs. In this study, the Gaussian, Student-T, Frank, Clayton and Gumbel Copula as described in Chapter 3, are tested to find the best fitted theoretical Copula family for modelling $C(F_R(r), F_G(g))$ (note that only the radar/gauge positive pairs are considered, $R \geq 0.1$ [mm/hour] and $G \geq 0.1$ [mm/hour]).

The density of the empirical Copulas is shown in Figure 6.4, together with the estimated theoretical Gumbel Copula for the radar/gauge positive pairs exemplary for the station Garmisch-Partenkirchen at summer time, 2006 to 2007. It can be seen that the empirical Copula density is asymmetric with respect to the opposite diagonal of the unit square. The highest density can be found at the upper right corner with the second highest density at the lower left corner, indicating a strong upper/lower tail dependence.

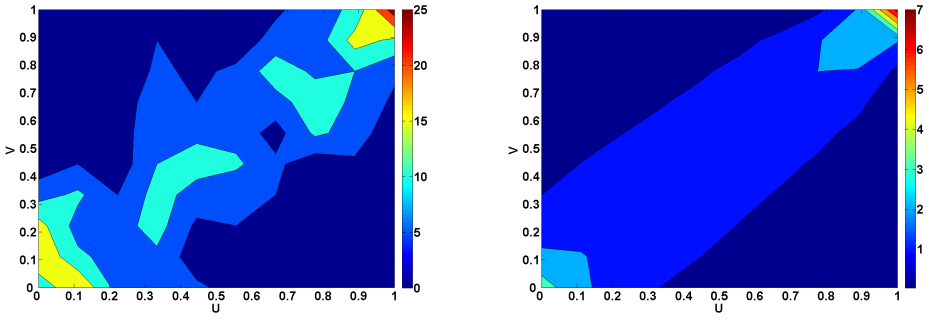

Figure 6.4: Density of the empirical and estimated Gumbel Copula (from left to right) calculated between radar/gauge positive pairs at Garmisch-Partenkirchen, summer, 2006 to 2007, w/o ARMA-GARCH transformation.

For the five different theoretical Copula functions, the S_n test is first applied, only to exclude the Gaussian and Student-T Copula. However, for the three Archimedean Copula, it is not sensitive enough. Therefore, the K-function test is further used to select the best fit one from Clayton, Frank and Gumbel Copula. The results of the K-function test are listed in the Table 6.2 for locations at Garmisch-Partenkirchen, Oberammergau and Wielenbach, summer time 2006 to 2007, as well as the estimated Copula parameters (MLE). The smallest test result means the best fitting and the minimum values are highlighted in Table 6.2.

It can be concluded that the Gumbel Copula show the best fitting, according to

Table 6.2: Goodness of Fit (GoF) test using the K-function. The minimum values of the K function value are highlighted in bold, suggesting the best fit, summer, 2006 to 2007, w/o ARMA-GARCH transformation.

Station	Gumbel		Frank		Clayton	
	K	θ_{G}	K	θ_{F}	K	θ_{C}
Garmisch-P.	**0.00078**	2.09	0.00092	6.77	0.003	2.31
Oberammergau	**0.00064**	2.44	0.001	8.37	0.003	2.85
Wielenbach	**0.000083**	2.25	0.001	7.98	0.004	2.77

the results from the GoF test. Furthermore, a strong asymmetric between upper and lower tail dependence can also be found in theoretical Gumbel PDF similar as in the empirical Copula PDF as shown in Figure 6.4, which can not be found in Gaussian and Student-T Copula with the symmetric upper and lower tail dependence structures. Therefore, in this thesis, the Gumbel Copula is chosen to model the dependence structure between radar and gauge positive pairs (the same for MW-link/radar and MW-link/gauge positive pairs also according to the GoF tests), for the case w/o ARMA-GARCH transformation.

Similar as the parameters of marginal distributions, the seasonal pattern can also be found in the fitted Copula parameters in different seasons as shown in Figure 6.5. Generally, the highest dependence between radar and gauge is reached in summer while the lowest in winter, also implying that the temperature can have impacts on the dependence structure between radar and gauge. For instance, snow may reduce the radar/gauge dependence structure during winter time because of the strong bright band effects. On the other hand, with the increasing of temperature, the radar/gauge dependence is also increasing. Another point worth to note is that, at the station Wielenbach located in the flat area, the highest dependence is not in summer but in autumn. Further discussion about this temperature dependent Copula can be found in Chapter 9.

6.3.2 Simulation and Validation

With the calculated marginal distributions and dependence function $c_\theta(F_R(r), F_G(g))$, random samples of gauge or called pseudo observations can be generated through the

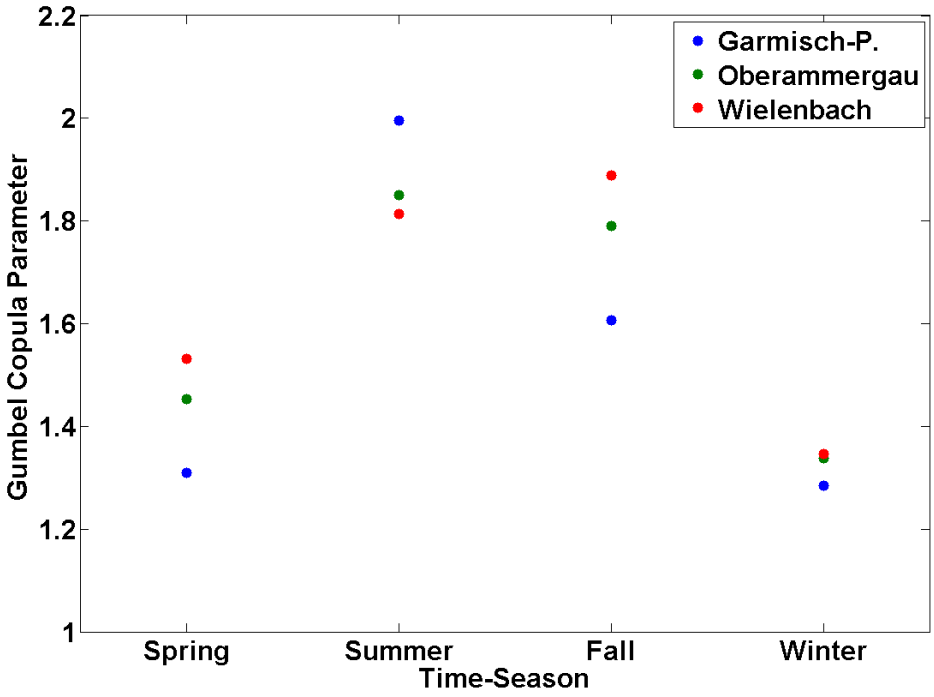

Figure 6.5: Estimated Copula parameter (Gumbel) of radar/gauge positive pairs at Garmisch-Partenkirchen, Oberammergau and Wielenbach, in different seasons, 2005 to 2008, w/o ARMA-GARCH transformation.

corresponding PDF conditioned on the given radar measurements. The details of this Copula-based simulation approach have already been described in Chapter 3.

One example of simulated pseudo observations at Garmisch-Partenkirchen is shown in the Figure 6.6. At the top, the box-plot can be considered as a simulation interval for each time step. This is due to the reason that a random sample with 100 values is generated from the corresponding conditional PDF based on the given radar measurement. From this point of view, not only one simulated pseudo observation value but also a full distribution is given by this Copula-based approach.

At the bottom of Figure 6.6, the box-plot is replaced by the mean of 100 randomly sampled values. A comparison with the gauge observations by using correlation, RMSE and NSE, shows that the mean value performs better than the median of 100 randomly sampled values. So, throughout this study, the pseudo observation always refers to the mean of the randomly samples (or equal to the conditional expectation value).

From these two example figures, it can also be seen that the over-estimation of radar measurements can be reduced in a reasonable way comparing to the gauge observations. Or in other word, the bias between radar and gauge can be corrected successfully.

Here, the pseudo observations are based on the radar measurements derived from the simple and standard Z/R (256/1.42) relationship as described in Chapter 4. In order to check the quality of this point wise Copula-based simulation technique, a more sophisticated algorithm to calculate radar observed rainfall - Radolan (DWD, 2001)- is further used. So, the pseudo observations, standard Z/R and Radolan derived radar observations are all compared to the gauge observations to check their performances, by using correlation coefficients, RMSE and NSE as described in Chapter 3. These validation results are presented in Table 6.3 at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich City.

Table 6.3: Comparison of the rainfall calculated with the simple Z/R (256/1.42) relationship, Radolan derived rainfall, Copula-based pseudo observations (both data and rank space) and rain gauges at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich City, only for positive pairs, summer, 2005 to 2008, w/o ARMA-GARCH transformation. The best ones are highlighted in bold.

Station	Garmisch-P.			Oberammergau		
Methods	Corr [–]	RMSE [mm/h]	NSE [–]	Corr [–]	RMSE [mm/h]	NSE [–]
Simple Z/R	0.64	1.74	0.21	0.60	1.82	0.19
Radolan	**0.66**	1.66	0.28	**0.63**	1.71	0.28
Pseudo.O	0.65	**1.50**	**0.41**	0.63	**1.60**	**0.37**
Pseudo.O(rank)	0.70	-	-	0.66	-	-

Station	Wielenbach			Munich city		
Methods	Corr [–]	RMSE [mm/h]	NSE [–]	Corr [–]	RMSE [mm/h]	NSE [–]
Simple Z/R	0.53	1.72	0.20	0.57	1.81	0.16
Radolan	**0.58**	1.64	0.27	**0.55**	1.95	0.02
Pseudo.O	0.57	**1.59**	**0.32**	0.58	**1.61**	**0.33**
Pseudo.O(rank)	0.65	-	-	0.60	-	-

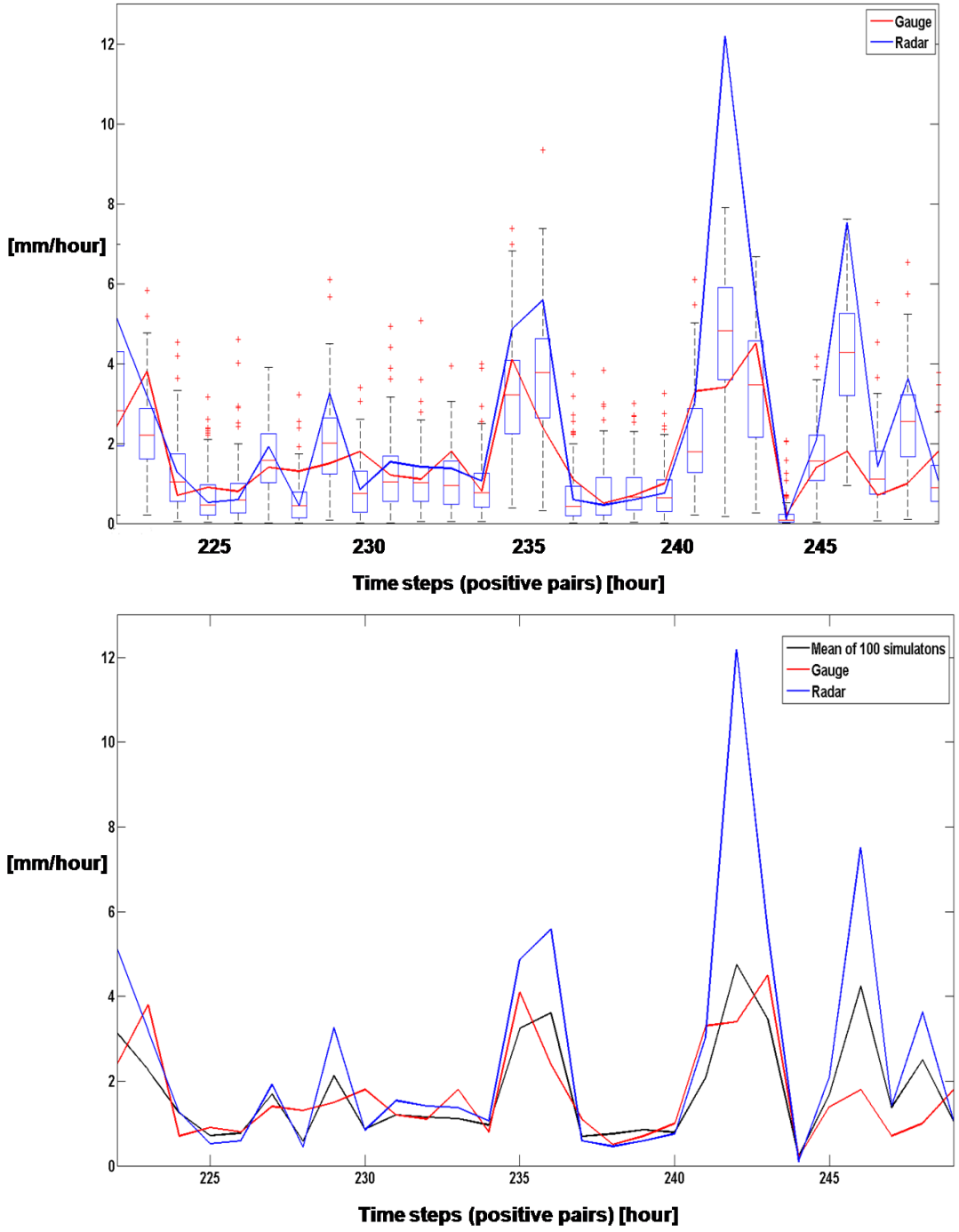

Figure 6.6: Time series of positive pairs of radar (Z/R-256/1.42) and gauge at Garmisch-Partenkirchen and box-plot/mean (top/bottom) of the random sample (100 realizations) of pseudo observations generated by using the Gumbel Copula, summer 2008, w/o ARMA-GARCH transformation.

According to the validation results, the pseudo observations (either in data space or rank space) have the best match with the gauge observations compared to the other two radar derived precipitation time series. For correlation, the pseudo observations in data space is not better than the Radolan results. However, the RMSE and NSE show that the pseudo observation improve the rainfall estimation, around 20 percent better in NSE compared to the Radolan and even more compared to the simple Z/R. This indicates that the simulated pseudo observations perform well and can efficiently correct the bias in the radar measurements.

However, the pseudo observation in rank space has better results than those in data space for Pearson's correlation. The reason can be due to the back-transformation from the rank space to the data space by using the estimated rain gauge marginal distribution, which can also be found in Figure 6.1. The CDF value is almost 1 when the corresponding gauge rainfall is just around 6 mm/hour. As a result, there are serious impacts on the back-transformations from the rank to data space, resulting at the low rainfall value at data space even though the corresponding rank value is high, for instance the simulated rank values 0.8, 0.9 and 0.95 can lead to nearly the same value in the data space.

As shown in Figure 6.7, those red cycles with the y axis value around 7 mm/hour (in the range of the yellow bracket) have larger values in the rank space. This means that those red cycles should be with the value around 12 mm/hour in data space (in the positions where the yellow arrow is referring to) if the gauge marginal could perform well enough for the extreme large values (both in rank and data spaces). As shown in the top of Figure 6.4, the upper tail dependence structure is very strong for the gauge/radar empirical Copula which means gauge and radar observations are better correlated when they have the extreme large precipitation values. However, after back-transformation, the values in the data space are reduced remarkably leading to an obvious bias.

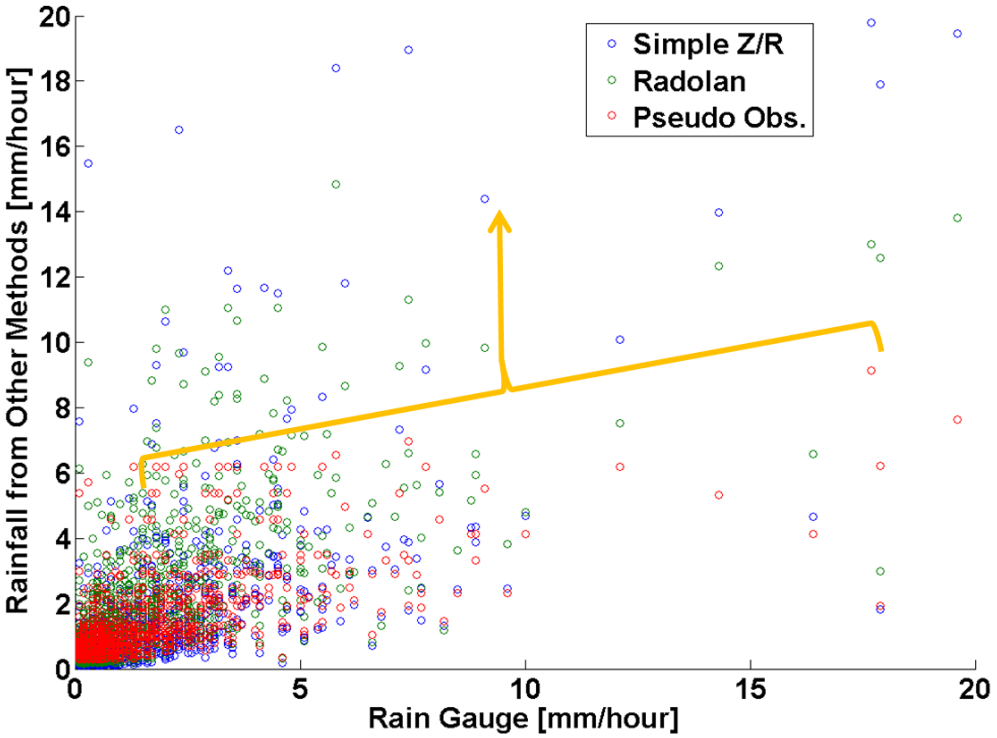

Figure 6.7: Scatter plot of radar derived rainfall (blue for simple Z/R, green for Radolan) and pseudo observations (red) at Garmisch-Partenkirchen, positive pairs, summer, 2005 to 2008, w/o ARMA-GARCH transformation.

6.4 Copula Analysis with ARMA-GARCH Transformation

6.4.1 Copula Fitting

In this approach including the ARMA-GARCH transformation, the marginal distributions are kept as the same as described in 6.2. The only difference comparing to the approach in 6.3 is that the dependence structure between radar and gauge positive pairs has to be investigated for the residuals after performing the ARMA-GARCH transformation. The densities of the new empirical Copula are shown in Figure 6.8, together with the estimated theoretical Gumbel Copula of the radar/gauge positive pairs at Garmisch-Partenkirchen, summer, 2006 to 2007. Similar as the results w/o ARMA-GARCH transformation, the upper/lower tail dependence can still be found,

but not so strong as in Figure 6.4.

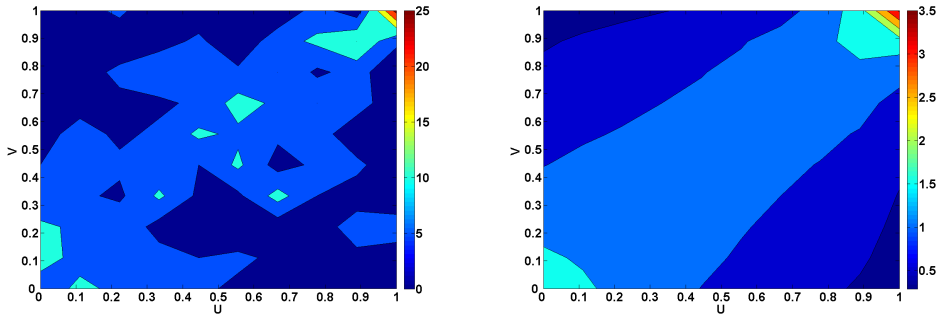

Figure 6.8: Density of the empirical and estimated Gumbel Copula (from left to right) estimated between radar and gauge positive pairs at Garmisch-Partenkirchen, summer, 2006 to 2007, w/ ARMA-GARCH transformation.

Similar as the approach w/o ARMA-GARCH transformation, the Gaussian and Student-T Copula can be excluded by using the S_n test. However, the K-function test is still needed to select the best fit one among the Clayton, Frank and Gumbel Copula. According to the K-function test results as listed in Table 6.4, the Gumbel Copula shows the best fit. Therefore, in this thesis, all the results w/ ARMA-GARCH transformation are generated by using Gumbel Copula. Compared to the results w/o ARMA-GARCH transformation, the main difference is that the new estimated Copula parameters are decreased to their 3/4 - 3/5, indicating a much more scattered dependence structure between gauge and the corresponding radar positive observations.

Table 6.4: Goodness of Fit (GoF) test using the K-function. The minimum values of the K function value are highlighted in bold, summer, 2006 to 2007, w/ ARMA-GARCH transformation.

Station	Gumbel		Frank		Clayton	
	K	θ_G	K	θ_F	K	θ_C
Garmisch-P.	**0.00013**	1.53	0.00052	3.50	0.0016	0.83
Oberammergau	**0.00022**	1.57	0.00077	3.57	0.0022	0.71
Wielenbach	**0.00013**	1.58	0.00054	3.71	0.0013	0.88

6.4.2 Simulation and Validation

The same as introduced in Chapter 6.3.2, with the estimation of marginal distributions and Copula parameters, the Copula-based simulation can be performed to generate the pseudo observations. One example of simulated pseudo observations at Garmisch-Partenkirchen is shown in the Figrue 6.9, both w/o and w/ ARMA-GARCH transformation (top to bottom). It can be found that:

1. The simulated pseudo observations w/ ARMA-GARCH are decreased when the conditioned radar values are large, compared to the results w/o ARMA-GARCH.

2. The simulated pseudo observations w/ ARMA-GARCH are increased when the conditioned radar values are small, compared to the results w/o ARMA-GARCH.

The simulated pseudo observations are compared to the gauge observations (summer, 2008) all the gauge stations and the results are listed in Table 6.5 and Table 6.6, both w/o and w/ ARMA-GARCH transformation. Note that the validation results for station Hohenpeissenberg are not presented as this rain gauge is located close to the radar station itself so that the measurements at this grid are obviously not qualified. It can be found that:

1. For Kendall's τ and Pearson correlation coefficient, both of the results are nearly the same although the simulations w/ ARMA-GARCH show a small decrease.

2. For RMSE, MAE, MSE and NSE, the difference is much larger. The results w/ ARMA-GARCH are 20 to 30 percent worse than those w/o ARMA-GARCH.

The reason for this is due to the reduced Copula parameters estimated from the original positive pairs w/o ARMA-GARCH and residuals w/ ARMA-GARCH.

6.5 Summary and Discussion

In this chapter, a point wise Copula-based data integration approach is introduced to combine precipitation information from radar and gauge. This method focuses on studying the dependence structure between gauge and radar positive pairs, using the Copula technique to model the joint distribution. Then, the pseudo observation can be generated from this joint distribution. By comparing the results of the Radolan derived

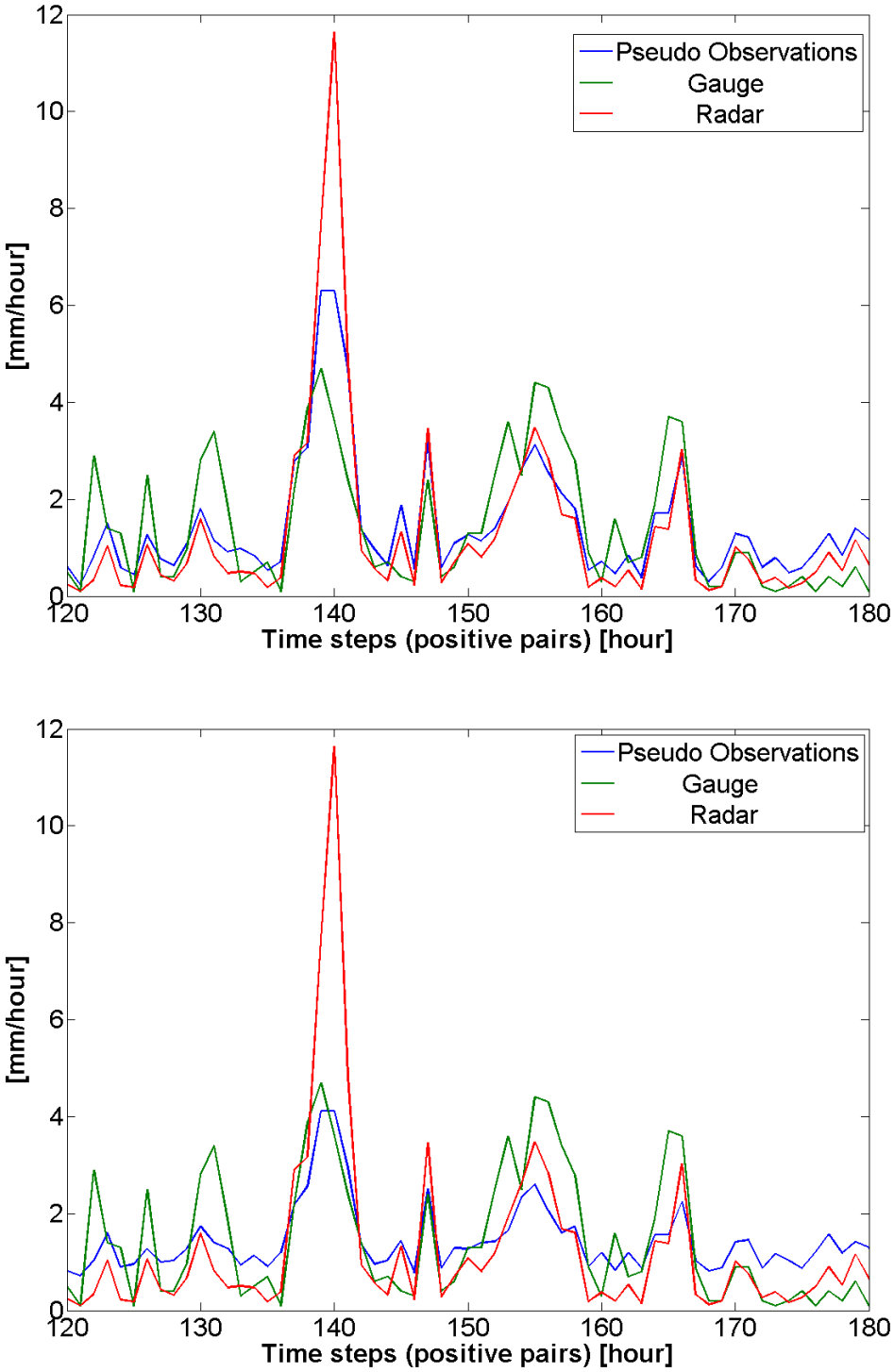

Figure 6.9: Time series of positive pairs of radar (Z/R-256/1.42), gauge and generated pseudo observations at Garmisch-Partenkirchen, summer 2008, w/o and w/ ARMA-GARCH transformation (top/bottom).

Table 6.5: Validation for Copula-based simulated pseudo observations for all stations, positive pairs only, summer, 2008, w/o ARMA-GARCH transformation.

.

ID	Station Name	Kendall's τ [–]	r [–]	RMSE [mm/h]	MAE [mm/h]	NSE [–]
1	Bernbeuren-Prachtsried	0.65	0.82	1.18	0.70	0.63
2	Diessen	0.58	0.79	1.36	0.74	0.53
3	Deisenhofen	0.59	0.74	1.58	0.86	0.45
4	Ettal	0.58	0.74	1.45	0.86	0.47
5	Garmisch-Partenkirchen	0.54	0.65	1.61	0.89	0.41
6	Gilching	0.55	0.75	1.58	0.80	0.47
7	Griesen	0.53	0.70	1.37	0.79	0.46
8	Halblech	0.55	0.74	1.57	0.86	0.49
9	Hindelang	0.51	0.63	1.78	1.10	0.40
10	Hohenpeissenberg	–	–	–	–	–
11	Kaufbeuren	0.61	0.78	1.01	0.66	0.60
12	Kochel	0.59	0.78	1.24	0.77	0.57
13	Kohlgrub, Bad	0.63	0.78	1.39	0.77	0.57
14	Kraftisried	0.57	0.70	1.57	0.97	0.48
15	Kreuth	0.50	0.67	1.74	1.00	0.42
16	Krün	0.53	0.72	1.80	0.94	0.44
17	Lenggries	0.53	0.68	1.66	0.95	0.43
18	Maisach	0.47	0.72	1.43	0.83	0.46
19	Marktoberdorf	0.63	0.79	1.58	0.89	0.49
20	München	0.54	0.70	1.29	0.85	0.49
21	Oberammergau	0.60	0.79	1.27	0.79	0.59
22	Oberschleissheim	0.54	0.78	1.56	0.81	0.51
23	Oy	0.53	0.63	1.55	1.00	0.39
24	Schwangau	0.55	0.67	1.41	0.85	0.43
25	Seeg	0.61	0.68	1.71	0.91	0.41
26	Schäftlarn	0.61	0.81	1.15	0.73	0.63
27	Steingaden	0.69	0.83	1.67	0.75	0.55
28	Schwaben	0.55	0.55	1.79	0.94	0.30
29	Schlehdorf	0.61	0.74	1.70	0.89	0.47
30	Vilgertshofen	0.59	0.70	1.64	0.82	0.44
31	Wielenbach	0.59	0.76	1.36	0.77	0.53

Table 6.6: Validation for Copula-based simulated pseudo observations for all stations, positive pairs only, summer, 2008, w/ ARMA-GARCH transformation.

ID	Station Name	Kendall's τ [–]	r [–]	RMSE [mm/h]	MAE [mm/h]	NSE [–]
1	Bernbeuren-Prachtsried	0.58	0.79	1.47	0.94	0.42
2	Diessen	0.57	0.78	1.52	0.86	0.40
3	Deisenhofen	0.56	0.73	1.70	0.96	0.37
4	Ettal	0.55	0.73	1.57	0.97	0.38
5	Garmisch-Partenkirchen	0.52	0.65	1.73	1.00	0.32
6	Gilching	0.52	0.76	1.70	0.89	0.39
7	Griesen	0.53	0.70	1.48	0.84	0.37
8	Halblech	0.52	0.72	1.74	1.00	0.37
9	Hindelang	0.49	0.63	1.88	1.20	0.34
10	Hohenpeissenberg	–	–	–	–	–
11	Kaufbeuren	0.59	0.77	1.12	0.77	0.50
12	Kochel	0.60	0.78	1.32	0.83	0.51
13	Kohlgrub, Bad	0.61	0.78	1.55	0.88	0.47
14	Kraftisried	0.53	0.70	1.73	1.10	0.36
15	Kreuth	0.49	0.67	1.84	1.10	0.35
16	Krün	0.51	0.73	1.88	1.10	0.39
17	Lenggries	0.51	0.66	1.78	1.04	0.34
18	Maisach	0.42	0.70	1.60	0.97	0.32
19	Marktoberdorf	0.58	0.78	1.79	1.07	0.34
20	München	0.49	0.66	1.42	1.01	0.38
21	Oberammergau	0.58	0.77	1.42	0.90	0.49
22	Oberschleissheim	0.49	0.77	1.75	0.95	0.38
23	Oy	0.51	0.60	1.63	1.09	0.33
24	Schwangau	0.51	0.64	1.56	0.99	0.31
25	Seeg	0.56	0.68	1.84	1.02	0.32
26	Schäftlarn	0.58	0.79	1.35	0.94	0.49
27	Steingaden	0.66	0.84	1.87	0.92	0.43
28	Schwaben	0.55	0.55	1.81	0.99	0.29
29	Schlehdorf	0.60	0.75	1.80	0.99	0.41
30	Vilgertshofen	0.58	0.70	1.72	0.89	0.38
31	Wielenbach	0.57	0.76	1.56	0.90	0.38

rainfall and the pseudo observations with the real gauge observations, the Copula-based approach leads to a better rainfall estimation by reducing the bias between radar and gauge, at least for the case w/o ARMA-GARCH transformation. As a result, this point wise Copula-based integration method is a useful and efficient tool to combine the gauge and radar data.

In order to apply Copula-based analysis, it has to be checked carefully whether the data is *i.i.d.* or not. In this dissertation, both approaches, w/o or w/ ARMA-GARCH transformation, are applied to test the sensitivity. Although the simulation results w/o ARMA-GARCH are better in the sense of cross-validations, from theoretical point of view, the ARMA-GARCH transformation is still required to make the whole process more systematic. In the following chapters, the sensitivity of ARMA-GARCH transformation to the final precipitation fields will be further discussed.

Additionally, the seasonal behaviour, both for the marginal distributions and estimated Copula parameter, are investigated in this chapter. This is also the influence from the changing temperature in different seasons, indicating the impacts of energy fluxes and balance on the precipitation. This will be discussed in detail in the following chapters.

However, some other problems still remain. For instance, the impacts from gauge marginal distribution during the back-transformation from rank to data space. Better marginal distributions are required especially with better behaviour at regions with high precipitation values. For modelling the radar/gauge dependence structure, if the rainfall type classification could be done, then different Copula dependence functions can be calculated for different rainfall types. An improvement in the simulation results can be expected.

In the following chapter, the Copula-based spatial data assimilation approaches are developed based on this point wise approach.

7

Copula Parameter Map Based Approach to Assimilate Precipitation Information from Radar and Gauge

In the past, many efforts have been put on estimating rainfall fields by combining data from the point accurate gauge and spatial priority radar. The detailed introduction can be found in Chapter 1 and will not be repeated here. In this study, two Copula-based approaches are developed to assimilate precipitation information from gauge and radar. The basic principles for the Copula-based spatial data integration methods are listed as following:

1. The spatial rainfall patterns are retrieved from the radar observed fields.

2. The absolute values at specific location are chosen to rely on the rain gauges.

As a result, the spatial distribution of radar/gauge dependence structure is one of the two crucial aspects for the Copula-based rainfall field estimation, which are mainly introduced in this and the next chapter. The other important aspect, the spatial distribution of the marginal distributions for rain gauges, is described in detail in Chapter 9.

In this chapter, the first Copula-based approach is developed for assimilating information from radar fields and gauge by studying the dependence structure between

them. This is an extension from point wise data integration to the spatial case. Pseudo observation fields are simulated and then compared to the real gauge observations, interpolated gauge filed. The strength and weakness of this new method, in both cases of w/o and w/ ARMA-GARCH transformation, are presented and discussed according to the validation results.

This chapter is structured as following. First, in section 7.1, an introduction to the Copula parameter map based approach is given. Then two strategies, to make use of the Copula parameter maps for the purpose of spatial simulation, are described and applied in section 7.2, namely the *Multiple Theta* and the *Maximum Theta* method. Afterwards, the validations for those two approaches are presented in section 7.3. The simulation w/ ARMA-GARCH transformation is presented in section 7.4. Finally, in section 7.5, the summary and discussion are given.

7.1 Copula Parameter Map

As already presented in the last chapter, for those grids where both radar and gauge measurements exist, the marginal distributions are estimated and a theoretical Copula model is estimated to describe the dependence structure between gauge and corresponding radar. It is worth to note that the Copula parameter (here, referring to Gumbel's θ_{G}, the same in this thesis except for special note) is a measure for the strength of the dependency as it is directly linked to Kendall's τ. Larger Copula parameters indicate stronger correlation. In previous chapter, the calculated Copula parameters are restricted to those certain grids both with radar and gauge. However, the procedure can be repeated to find the appropriate θ between each grid in the domain and a certain rain gauge.

Preliminary, it is straightforward to assume that the Copula parameter θ is a function of distance d between the specific rain gauge and radar pixels, as shown below, implying the isotropic assumption.

$$\theta = f(d) \tag{7.1}$$

For example, by calculating the dependence structures between one station and all grid cells, then assigned to different distances between them, an empirical curve exemplary for the rain gauge at Wielenbach is presented in Figure 7.1. Generally, it

can be seen that the dependence decreases with increasing distance from this rain gauge. However, the fluctuation (variability) increases with increasing distance from the gauge station. Therefore, the isotropic assumption for the dependency is not reasonable here.

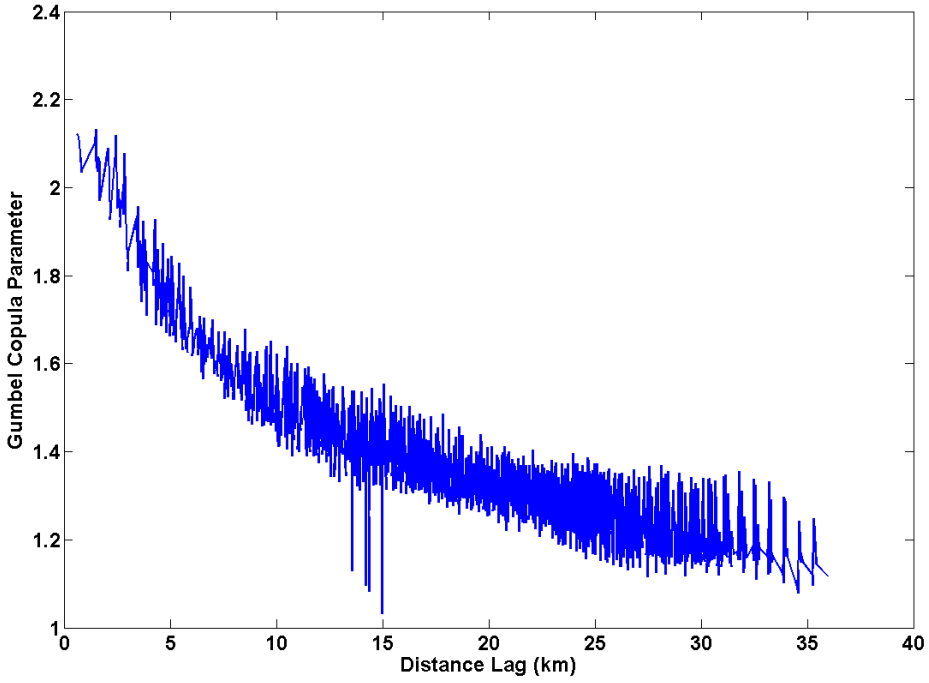

Figure 7.1: Gumbel Copula parameter θ_G as a function of radius, calculated between the gauge at Wielenbach and all the nearby radar grid cells, positive pairs only, summer, 2005 to 2008, w/o ARMA-GARCH transformation.

As a result, in addition to distance, the direction should also be considered, leading to the more realistic assumption - anisotropy for the dependence structure between all gauges and the radar field. The result, the so called Copula parameter map, is the reflection of the spatial dependence structure between one certain gauge and all radar grids in the whole domain. The Copula parameter maps are the basis for the first Copula-based spatial data integration approach.

The basic steps to produce the Copula parameter map are summarized as following:

1. For all rain gauges $G_i, i = 1, ..., n$ in the study area and all the radar pixels $R_j, j = 1, ..., m$, during certain training period, the corresponding marginal distributions

are estimated.

2. For one specific rain gauge, as well as all the m radar pixels, the corresponding Copula parameters θ_G for these m radar/gauge pairs are calculated.

3. The procedure is repeated for each rain gauge in the whole domain which resulting in n Copula parameter maps.

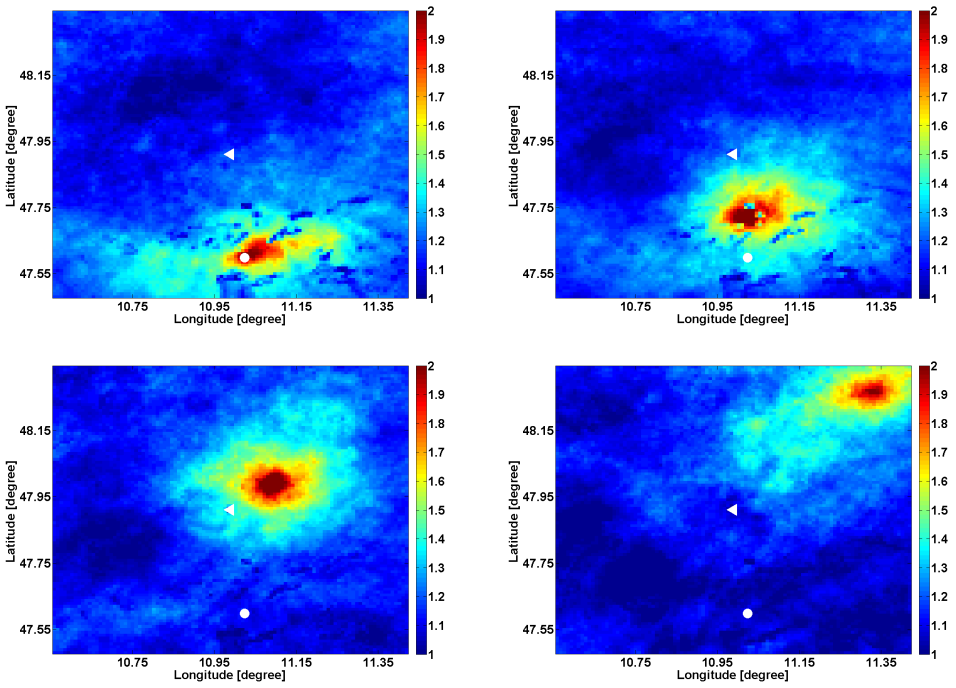

Figure 7.2: Copula parameter maps showing the Copula parameter θ_G calculated from the rain gauges at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich (from left to right and top to bottom), summer, 2007 to 2008, w/o ARMA-GARCH transformation. The white triangle refers to the radar station at Hohenpeissenberg and the white cycle refers to Garmisch-Partenkirchen, the same all through the thesis.

In the very idealised set-up where the assumption of isotropy can be applied, one would see a perfect rotationally symmetric parameter map, where the values decrease linearly with the distance from the gauge, as the Copula parameter map shown in Figure 7.2 derived from the station at Wielenbach. That is because Wielenbach is located at a flat area where the inhomogeneities introduced by e.g. systematic measurement errors, prevailing wind directions or complex terrain are relatively small. In contrast, the

Copula parameter map calculated from Garmisch-Partenkirchen, surrounded by a more complex terrain, shows significant asymmetries. This strong asymmetry is partly due to the reason that the Wetterstein and the Ammergauer Alps favour a westerly flow and introduce variability to the field due to strong height gradients, as similar asymmetric structure can also be found in maps derived from gauges in Oberammergau and Munich located at Pre-Alpine and city regions respectively.

It is concluded that the Copula parameter map is a reasonable way to represent the spatial dependence structures between radar and gauge, which is a purely statistical modelling algorithm and includes the information from terrain pattern naturally. The anisotropic nature of the dependence structures has to be considered when Copula-based precipitation fields are modelled. The two different Copula parameter map based approaches namely *Multiple Theta* and *Maximum Theta* are presented in the following sections.

7.2 Simulated Field of Pseudo Observations

Considering the spatial simulation of precipitation fields, the generated Copula parameter maps have to be combined in a reasonable way so that the statistical characteristics of all gauges can be assimilated simultaneously. Figure 7.3 schematically illustrates the two different methods developed to make use of the Copula parameter maps in this study, named *Multiple Theta* and *Maximum Theta* approaches, for one single time step.

7.2.1 Multiple Theta

The first possibility is to use all n theta maps (here $n = 31$) separately and create an ensemble of n samples of pseudo observations in the rank space, one for each map. A back-transformation by the probability integral transformation is performed by using the marginal distribution from where the used Copula parameter map is produced. These n back-transformed pseudo observations are combined by inverse distance weighting (IDW), which leads to one final result in the data space.

The basic steps for *Multiple Theta* approach can be summarized as follows:

1. Based on the available gauge stations n Copula parameter maps are derived, resulting in a set of n Copula parameters for a specific radar grid cell.

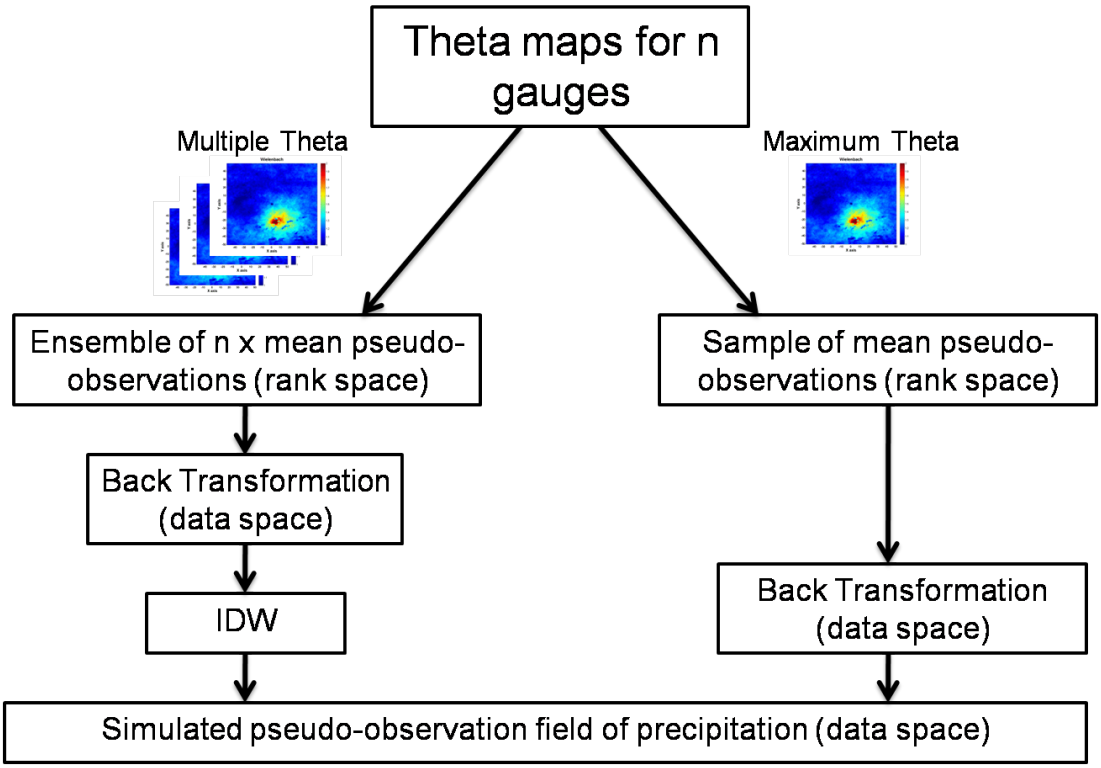

Figure 7.3: Flowchart of *Multiple Theta* and *Maximum Theta* approaches for Copula parameter map based rainfall field simulation.

2. n Copula parameters, n marginal distributions of the gauge stations and one single marginal distribution of the specific radar grid cell are assigned to a specific radar grid cell.

3. For each set (i.e. marginal distribution gauge, marginal distribution radar, and Copula parameter) a sample of 100 members is simulated in the rank space.

4. The expectation values are calculated from the random samples.

5. The integral transformation is applied to the calculated expectation values to transform back to data space.

6. These steps are repeated for all radar grid cells.

7. Finally, Inverse Distance Weighting (IDW) is applied to generate one single value for each radar grid cell.

This approach leads to a field of pseudo observations in the data space. At one time step of the test period in summer, 2008, by applying this *Multiple Theta* simulation approach, Figure 7.4 compares one example pseudo observation field and the corresponding radar field. It can be seen that, compared simulated field of pseudo observations to the original radar image, the basic spatial rainfall pattern is remained and the spokes and backscattering effects are reduced. For specific grid cells, the absolute values are corrected towards the ground measurements. The spatial variances of rainfall intensities are also reduced. Additionally, in Figure 7.5, the simulated fields of pseudo observation are presented for the complete rainfall event from 08:00 to 16:00, 04.08.2008.

7.2.2 Maximum Theta

The so called *Maximum Theta* approach combines all the available Copula parameter maps by comparing the n calculated Copula parameters from all available Copula parameter maps for each grid cell and choosing the maximum value. However, this can only be done when the Copula has only one parameter. So, the *Maximum Theta* approach can not be generally applicable for all Copula families, not like the *Multiple Theta* approach.

The back-transformation in this approach uses the marginal distribution of the gauge that is corresponding to the maximum Copula parameter, making sure that the method is coherent. As each of the maps contains certain asymmetries, the maximum theta map will show very special structures. The advantage compared to the inverse distance weighting procedure is that the influence of each station can be determined according to the asymmetries. This means that any effects regarding correlations between one single radar pixel and a certain gauge such as complex terrain, dominant flow directions and many others will be respected. As a consequence not the gauge which lies closest to a certain radar pixel will dominate the process but the most correlated one will be regarded most. Eventually, the basic steps for multiple theta simulation can be summarized as follows:

1. Based on the available gauge stations n Copula parameter maps are derived, resulting in a set of n Copula parameters for a specific radar grid cell.

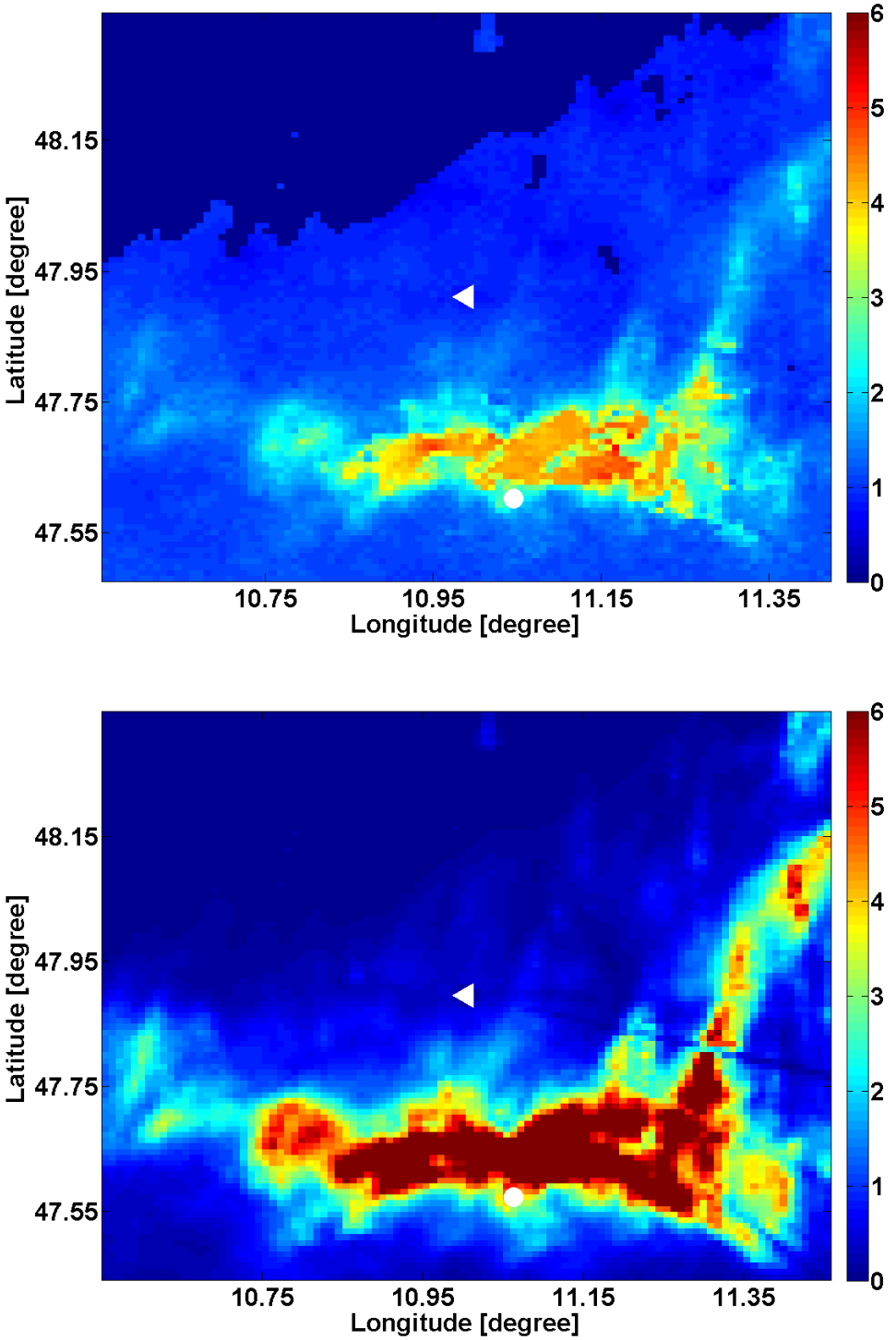

Figure 7.4: Simulated field of pseudo observations derived by the *Multiple Theta* approach (top) and uncorrected radar field (bottom), 13:00, 04.08.2008, Gumbel Copula, w/o ARMA-GARCH transformation. The unit of the color bar is in [mm/hour].

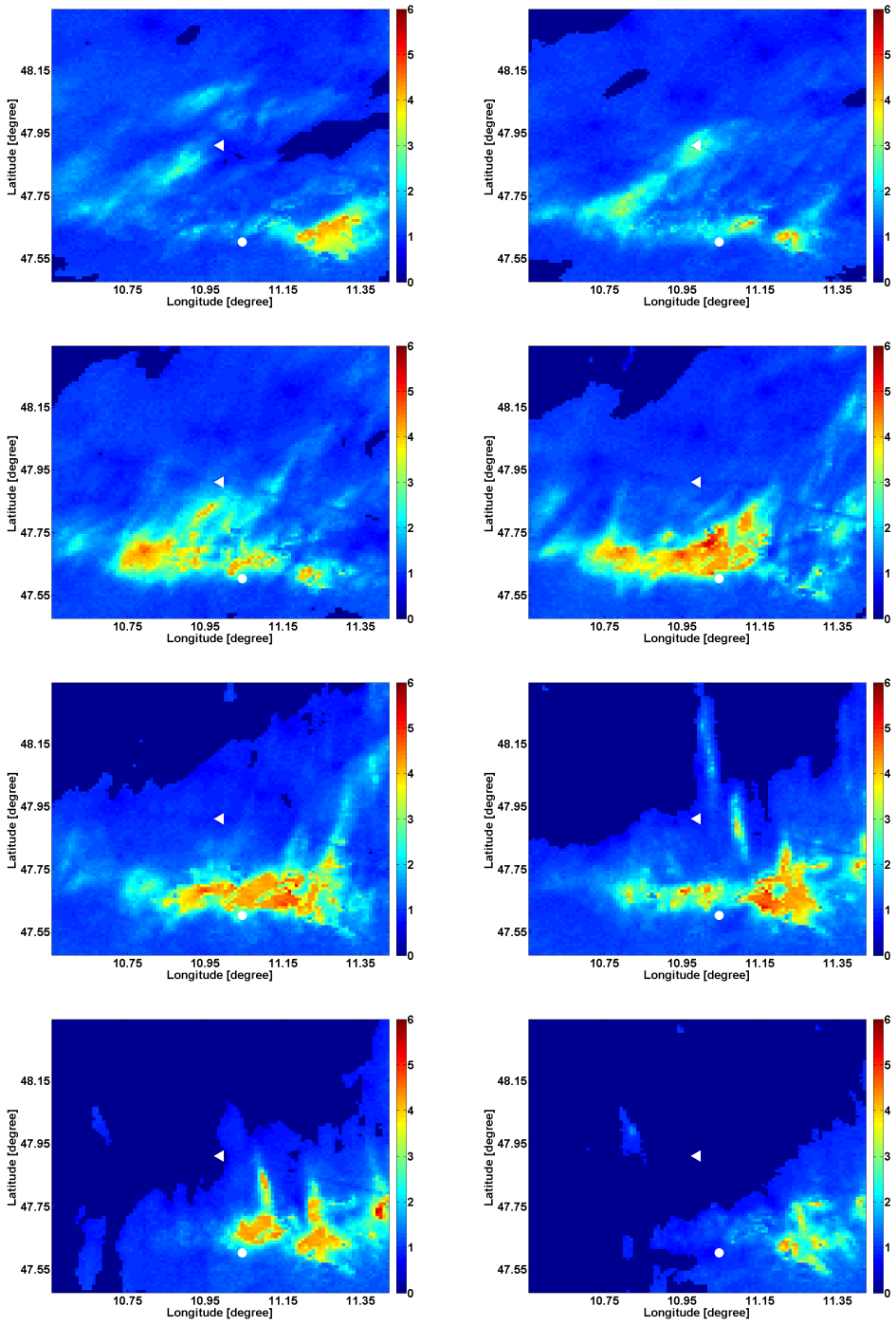

Figure 7.5: Simulated fields of pseudo observations derived by the *Multiple Theta* approach for a complete rainfall event from 09:00 to 16:00 (from left to right and top to bottom), 04.08.2008, Gumbel Copula, w/o ARMA-GARCH transformation. The unit of the color bar is in [mm/hour].

2. The set showing the maximum Copula parameter is assigned for a specific grid cell, retaining the information of the Copula parameter and the corresponding marginal distributions (both gauge and radar).

3. One sample of 100 members is simulated in the rank space for this set.

4. The expectation value is calculated from the random samples.

5. The integral transformation is applied to the expectation values to transform back to data space.

6. These steps are repeated for all radar grid cells.

7. A field containing the expectation values for each radar grid cell is obtained.

Figure 7.6 shows the maximum theta map for the whole domain. The asymmetries from the single station maps are transported into the field, making sure that all possible variabilities in the dependence structures are considered.

At one time step of the test period - 13:00, 04.08.2008, by applying this *Maximum Theta* approach, Figure 7.7 compares one example field of pseudo observations and the corresponding radar field can be seen at the bottom of Figure 7.4. The *Maximum Theta* approach also has the ability to well reproduce the patterns from the radar observed rainfall field. Compared to the simulated field in the top of Figure 7.4 by using the *Maximum Theta* approach, the result field is not so smoothed with more details retained and the absolute values are higher. Additionally, in Figure 7.8, the simulated fields of pseudo observation are presented for the complete rainfall event from 08:00 to 16:00, 04.08.2008.

7.3 Validation

In order to check the quality of simulated pseudo observation field, different approaches are performed such as visual and quantitative validation by comparing to interpolated gauge rainfall field and real gauge measurements at grid scale. The validation results are presented in this section.

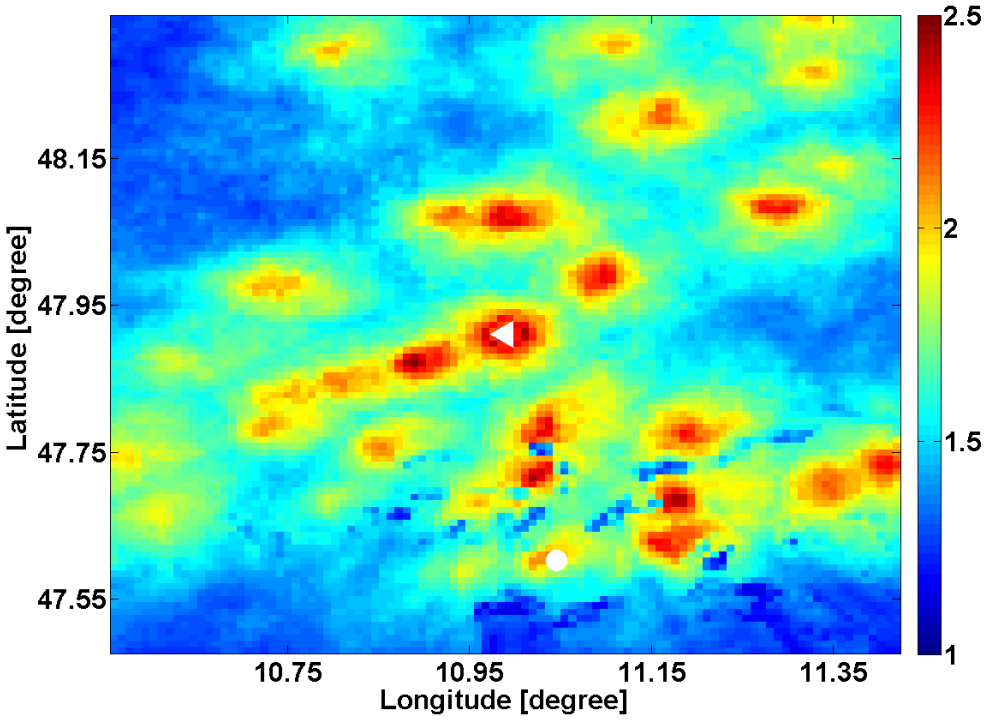

Figure 7.6: Maximum theta map derived from all available Copula parameter maps, summer, 2006 to 2007, Gumbel Copula, w/o ARMA-GARCH transformation. The color bar is the value of the Copula parameter θ_G .

7.3.1 Visual Inspection

By performing an Ordinary Kriging approach, the precipitation measurements from the 31 gauges in this study area were interpolated to derive the rainfall field as shown in Figure 7.9 at 13:00, 04.08.2008.

The spatial pattern of rainfall derived from the original radar field (see at the bottom of Fig.7.4) is roughly reproduced in the interpolated precipitation field (Fig.7.9). However, due to the limited number of gauges distributed in the whole domain (only covering the central region, see Fig.4.3), the precipitation field is smoothed remarkably and concentrated in three local maxima. Therefore, the spatial variability of the precipitation can not be reproduced well in the gauge sparse regions by using the Ordinary Kriging approach. In contrast, the precipitation fields derived from the Copula parameter map based approach (see the top of Fig.7.4 and Fig.7.7) reproduce the overall

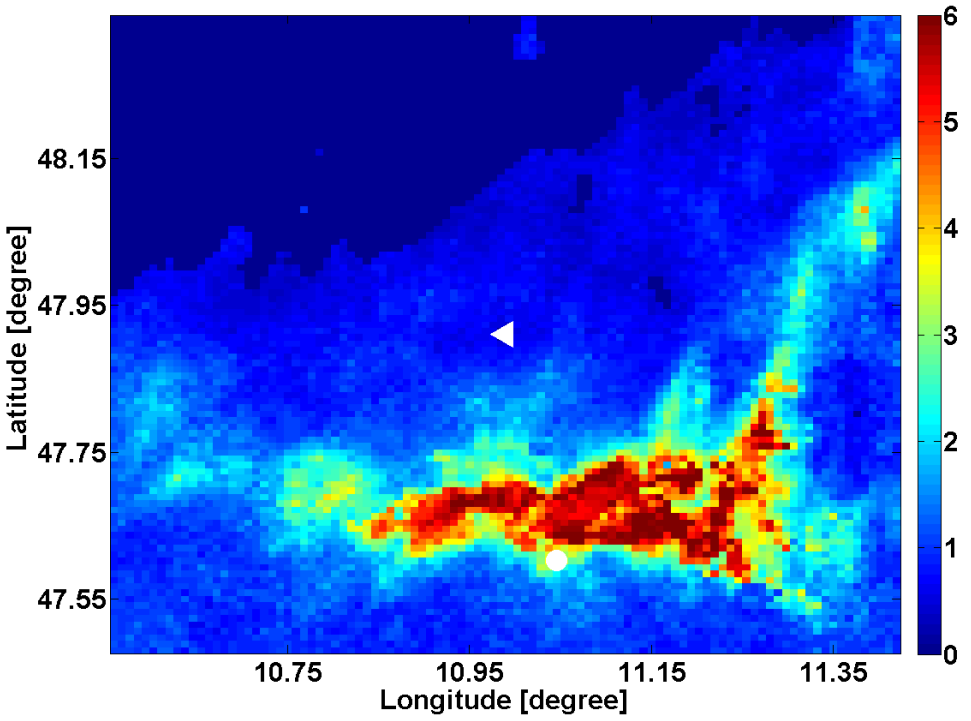

Figure 7.7: Simulated field of pseudo observations derived by the *Maximum Theta* approach, 13:00, 04.08.2008, Gumbel Copula, w/o ARMA-GARCH transformation. The unit of the color bar is in [mm/hour].

pattern of precipitation. This is because the precipitation information both from the radar and rain gauge are included by using the Copula parameter map based approach so that it is assumed to be superior to the traditional interpolation methods such as the Ordinary Kriging.

7.3.2 Quantitative Validation

The simulated time series of pseudo observation by using the *Maximum Theta* approach, as well as the corresponding gauge and radar observations, are presented in Figure 7.10 at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich during summer 2008, w/o ARMA-GARCH transformation. It can be seen that, the over-estimations (the peaks) in the radar observed rainfall are corrected by using the *Maximum Theta* approach. However, the under-estimations in the radar measurements are not well

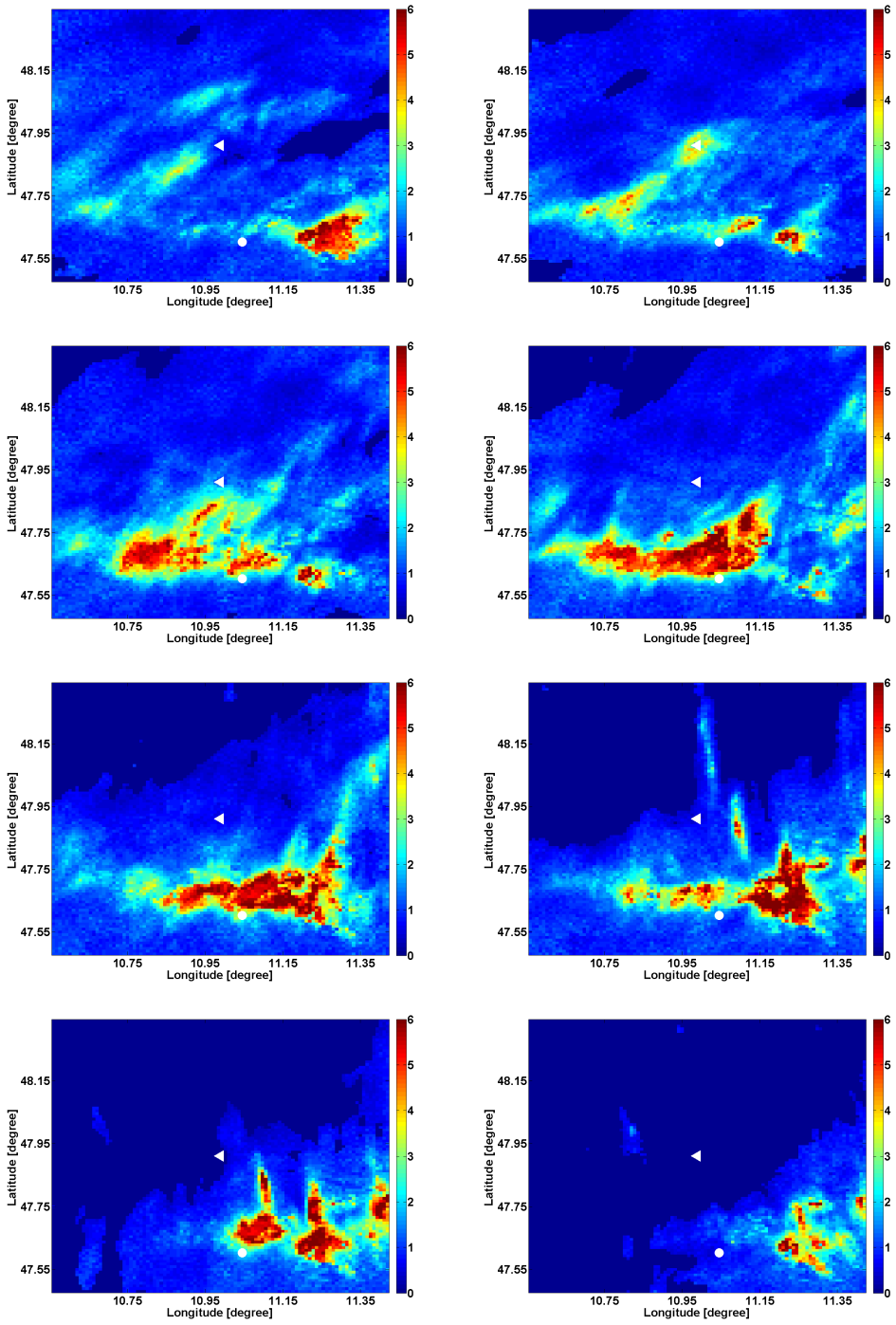

Figure 7.8: Simulated field of pseudo observations derived by the *Maximum Theta* approach for a complete rainfall event from 09:00 to 16:00 (from left to right, top to bottom), 04.08.2008, Gumbel Copula, w/o ARMA-GARCH transformation. The unit of the color bar is in [mm/hour].

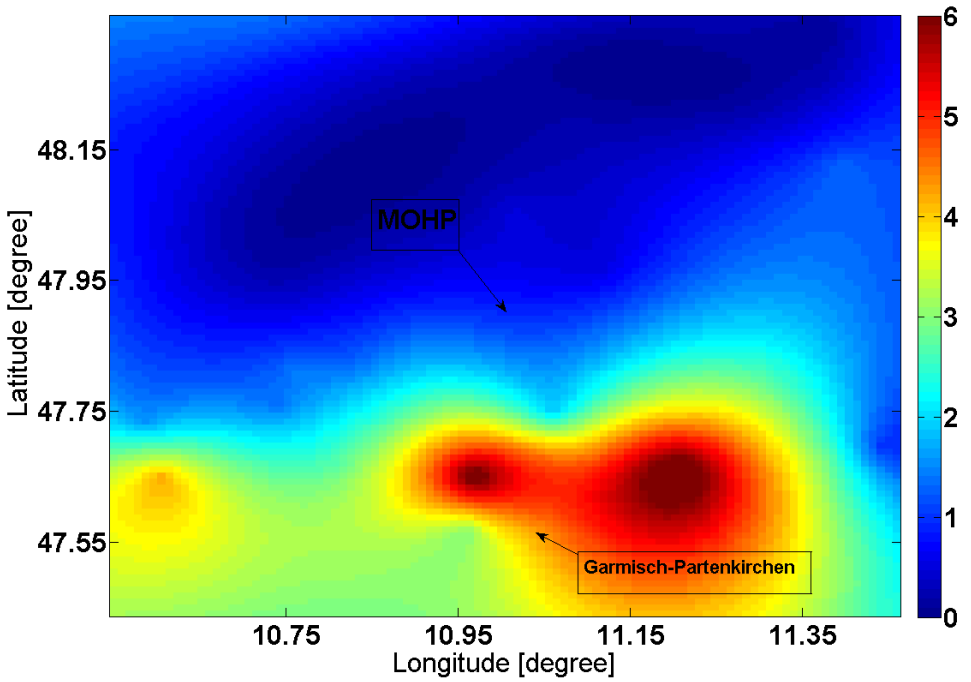

Figure 7.9: Interpolated rainfall field derived from the observed precipitation values at 13:00, 04.08.2008. The observations from the 31 gauge stations in the radar domain were interpolated by application of an Ordinary Kriging approach. The unit of the color bar is in [mm/hour].

corrected.

By applying point wise cross-validation, the pseudo observations generated from the *Maximum Theta* and *Multiple Theta* approaches are tested through the comparison with the real gauge time series in the respective grid cell, for all 30 gauges.

The point wise cross-validation results (see Table 7.2 and Table 7.1) by using different validation measures are used to test the performance of the simulation results quantitatively. In both tables, the mean value of Pearson's correlation coefficient is 0.7 calculated between the simulated positive series of pseudo observations and the real gauge observations, showing that the both proposed methods are equally good in this case. However, according to the results calculated by using RMSE and MAE, it is concluded that the *Maximum Theta* approach is slightly superior than the *Multiple Theta* approach. Similar results can also be found in the respective NSE values, on the mean

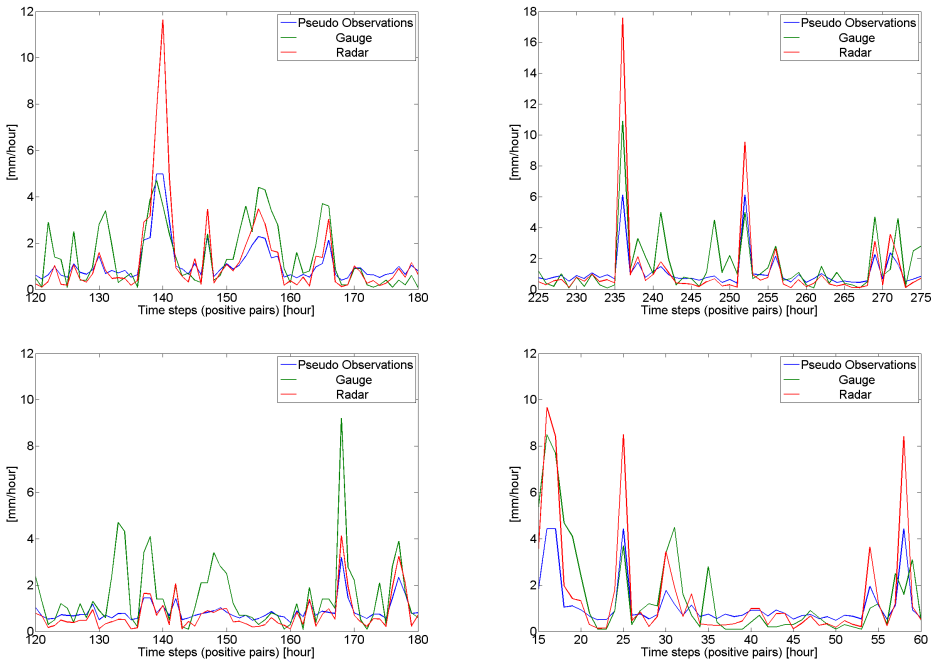

Figure 7.10: Time series of radar (Z/R-256/1.42), gauges and pseudo observations generated by using the *Maximum Theta* approach, positive pairs, at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich (from left to right and top to bottom), summer, 2008, Gumbel Copula, w/o ARMA-GARCH transformation.

level, 0.28 for the *Maximum Theta* approach and 0.19 for the *Multiple Theta* approach indicating 65 percent better. Note that the validation results for station Hohenpeissenberg are not presented as this station is located close to the radar observatory itself so that the measurements at this grid are obviously not qualified.

7.4 Simulations with ARMA-GARCH Transformation

The simulations in the previous parts are based on the assumption that both the radar/gauge observations (positive pairs) are already *i.i.d.*. If this constrain is not fulfilled, then following the algorithm introduced in (Laux et al., 2011) and also as described in Chapter 3 and 4, the ARMA-GARCH transformation should be applied first to remove the autocorrelation and heteroskedasticity structure in order to capture the real dependence structure between radar and gauge observations. The other simulation steps, for instance the estimation of marginal and simulation of pseudo observations,

Table 7.1: Point wise cross-validation for the *Multiple Theta* approach during 2008 summer, for all stations, positive pairs only, Gumbel Copula, w/o ARMA-GARCH transformation.

ID	Station Name	Kendall's τ [–]	r [–]	RMSE [mm/h]	MAE [mm/h]	NSE [–]
1	Bernbeuren-Prachtsried	0.65	0.77	1.75	1.03	0.19
2	Diessen	0.59	0.81	1.76	1.00	0.20
3	Deisenhofen	0.57	0.70	2.01	1.14	0.11
4	Ettal	0.56	0.68	1.76	1.09	0.23
5	Garmisch-Partenkirchen	0.55	0.64	1.82	1.02	0.24
6	Gilching	0.53	0.79	1.90	1.01	0.23
7	Griesen	0.53	0.71	1.55	0.89	0.31
8	Halblech	0.55	0.69	2.01	1.09	0.16
9	Hindelang	0.51	0.60	2.16	1.21	0.12
10	Hohenpeissenberg	–	–	–	–	–
11	Kaufbeuren	0.60	0.78	1.41	0.93	0.21
12	Kochel	0.60	0.74	1.67	0.99	0.22
13	Kohlgrub, Bad	0.62	0.80	1.91	1.08	0.19
14	Kraftisried	0.57	0.71	2.01	1.16	0.14
15	Kreuth	0.51	0.68	2.01	1.18	0.16
16	Krün	0.53	0.70	2.05	1.12	0.27
17	Lenggries	0.54	0.68	1.96	1.09	0.20
18	Maisach	0.47	0.75	1.72	1.02	0.22
19	Marktoberdorf	0.64	0.74	1.98	1.19	0.19
20	München	0.53	0.63	1.61	1.06	0.20
21	Oberammergau	0.60	0.76	1.72	0.96	0.25
22	Oberschleissheim	0.52	0.75	2.07	1.11	0.14
23	Oy	0.53	0.60	1.86	1.16	0.12
24	Schwangau	0.56	0.69	1.68	1.01	0.20
25	Seeg	0.60	0.67	2.08	1.14	0.14
26	Schäftlarn	0.60	0.80	1.76	1.04	0.14
27	Steingaden	0.70	0.90	2.14	1.09	0.25
28	Schwaben	0.54	0.49	2.02	1.14	0.10
29	Schlehdorf	0.60	0.73	2.16	1.20	0.14
30	Vilgertshofen	0.60	0.70	2.00	1.08	0.19
31	Wielenbach	0.58	0.71	1.74	0.97	0.23

Table 7.2: Point wise cross-validation for *Maximum Theta* approach during 2008 summer, for all stations, positive pairs only, Gumbel Copula, w/o ARMA-GARCH transformation.

ID	Station Name	Kendall's τ [–]	r [–]	RMSE [mm/h]	MAE [mm/h]	NSE [–]
1	Bernbeuren-Prachtsried	0.64	0.76	1.70	0.96	0.23
2	Diessen	0.52	0.81	1.68	0.88	0.27
3	Deisenhofen	0.55	0.69	1.84	0.99	0.26
4	Ettal	0.55	0.70	1.61	0.98	0.35
5	Garmisch-Partenkirchen	0.54	0.64	1.74	0.92	0.31
6	Gilching	0.49	0.79	1.78	0.90	0.33
7	Griesen	0.51	0.70	1.46	0.84	0.39
8	Halblech	0.51	0.69	1.93	0.98	0.23
9	Hindelang	0.50	0.59	2.02	1.09	0.23
10	Hohenpeissenberg	–	–	–	–	–
11	Kaufbeuren	0.56	0.75	1.23	0.78	0.40
12	Kochel	0.57	0.74	1.58	0.91	0.30
13	Kohlgrub, Bad	0.62	0.80	1.75	0.97	0.32
14	Kraftisried	0.49	0.69	1.93	1.08	0.21
15	Kreuth	0.52	0.68	1.95	1.11	0.27
16	Krün	0.51	0.69	1.92	1.02	0.36
17	Lenggries	0.53	0.68	1.74	0.96	0.38
18	Maisach	0.45	0.74	1.60	0.91	0.33
19	Marktoberdorf	0.62	0.75	1.86	1.08	0.29
20	München	0.52	0.63	1.57	0.95	0.24
21	Oberammergau	0.58	0.76	1.59	0.89	0.35
22	Oberschleissheim	0.51	0.71	1.93	0.96	0.25
23	Oy	0.53	0.61	1.68	1.03	0.29
24	Schwangau	0.53	0.61	1.68	0.90	0.31
25	Seeg	0.57	0.67	2.01	1.06	0.15
26	Schäftlarn	0.59	0.78	1.71	0.95	0.18
27	Steingaden	0.69	0.89	1.90	0.93	0.41
28	Schwaben	0.51	0.55	2.00	1.05	0.12
29	Schlehdorf	0.59	0.74	2.03	1.05	0.25
30	Vilgertshofen	0.59	0.70	1.84	0.95	0.29
31	Wielenbach	0.55	0.72	1.68	0.89	0.28

are kept as the same. The resulting Copula parameter maps are listed as in Figure 7.11, from the rain gauges at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich (Note that Gumbel Copula is used here according to the results from GoF tests).

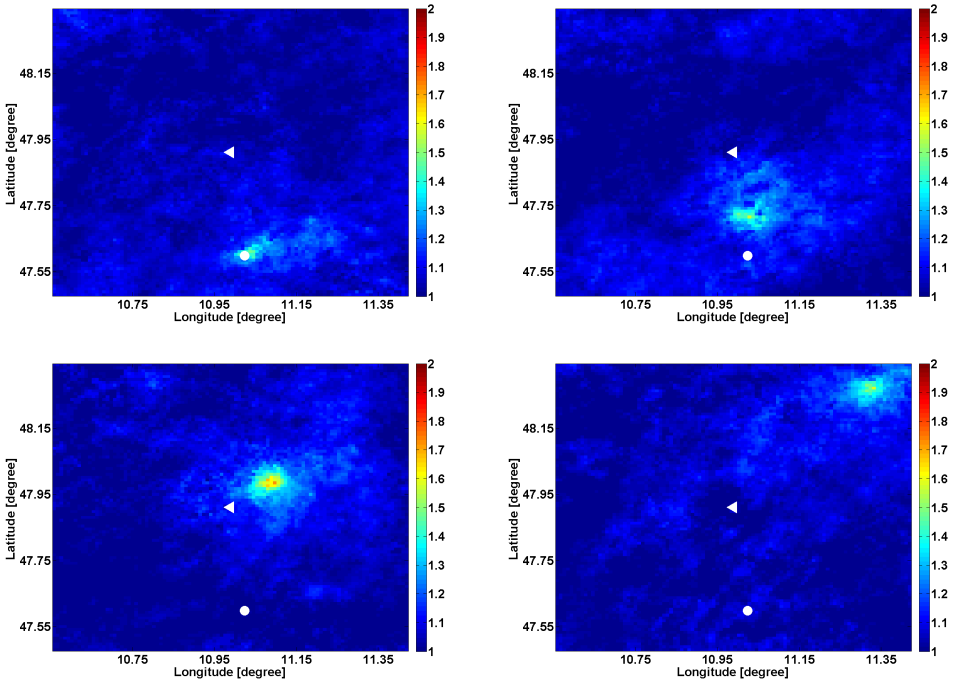

Figure 7.11: Copula parameter maps showing the Copula parameter θ_G calculated from the rain gauges at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich (from left to right and top to bottom), summer, 2007 to 2008, w/ ARMA-GARCH transformation. The color bar is the value of the Copula parameter θ_G.

The basic patterns of those Copula parameter maps are similar as shown in Figure 7.2 and Figure 7.11. The Copula parameter decreases with the increasing distance to the gauge and also indicates the asymmetric dependence structure influenced by the local topography and circulation pattern. However, the Copula parameters calculated from the *i.i.d.* residuals are reduced to around 3/4 of the values as presented in Figure 7.2 due to the impact from the ARMA-GARCH transformation.

Then based on these new Copula parameter maps, the *Maximum Theta* and *Multiple Theta* approaches, described in 7.2, can also be applied to generate the rainfall fields of pseudo observations and the examples of simulation results are presented in Figure

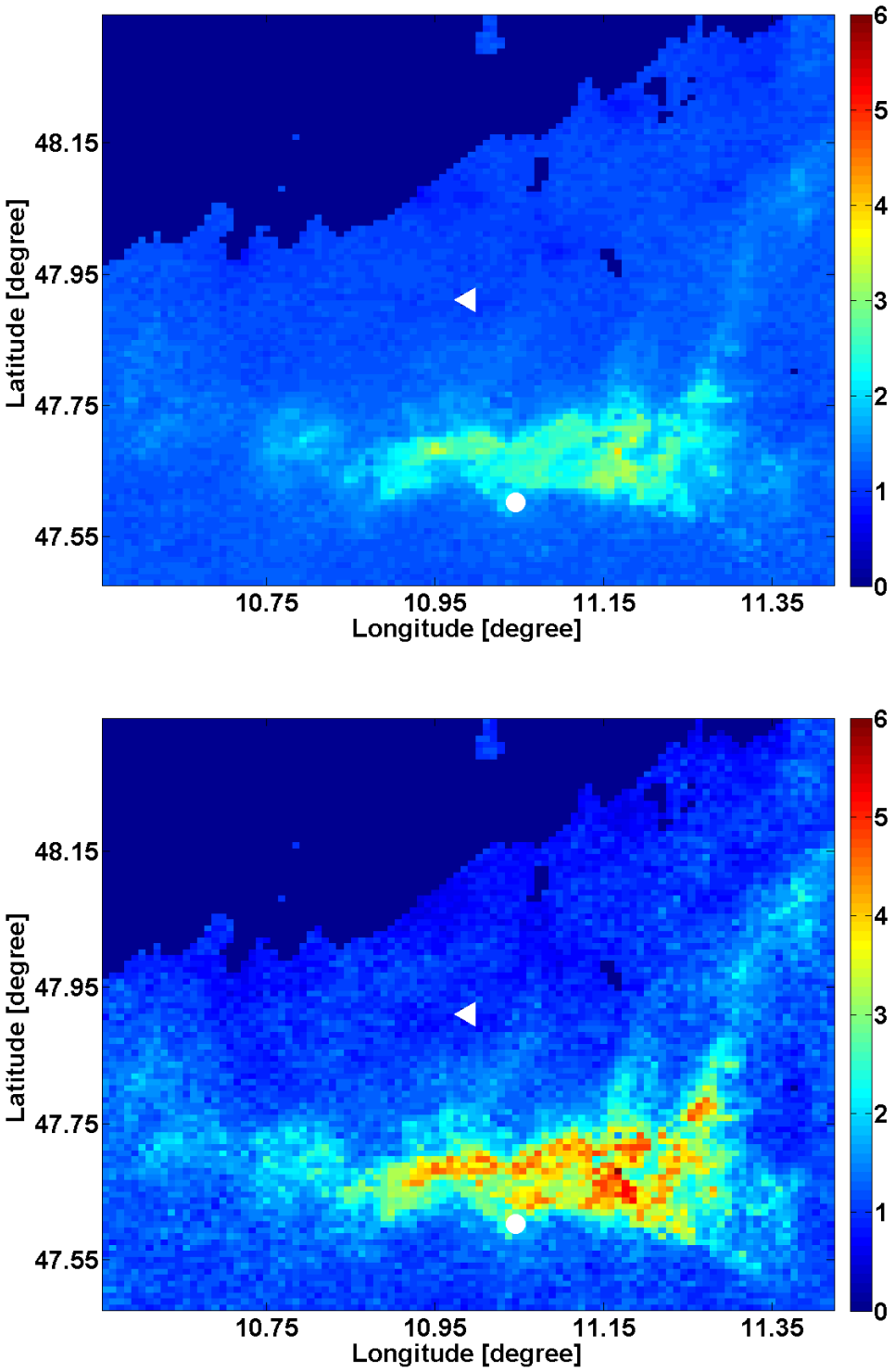

Figure 7.12: Simulated field of pseudo observations derived by the *Multiple Theta* and *Maximum Theta* approaches (from top to bottom), 13:00, 04.08.2008, Gumbel Copula, w/ ARMA-GARCH transformation. The unit of the color bar is in [mm/hour].

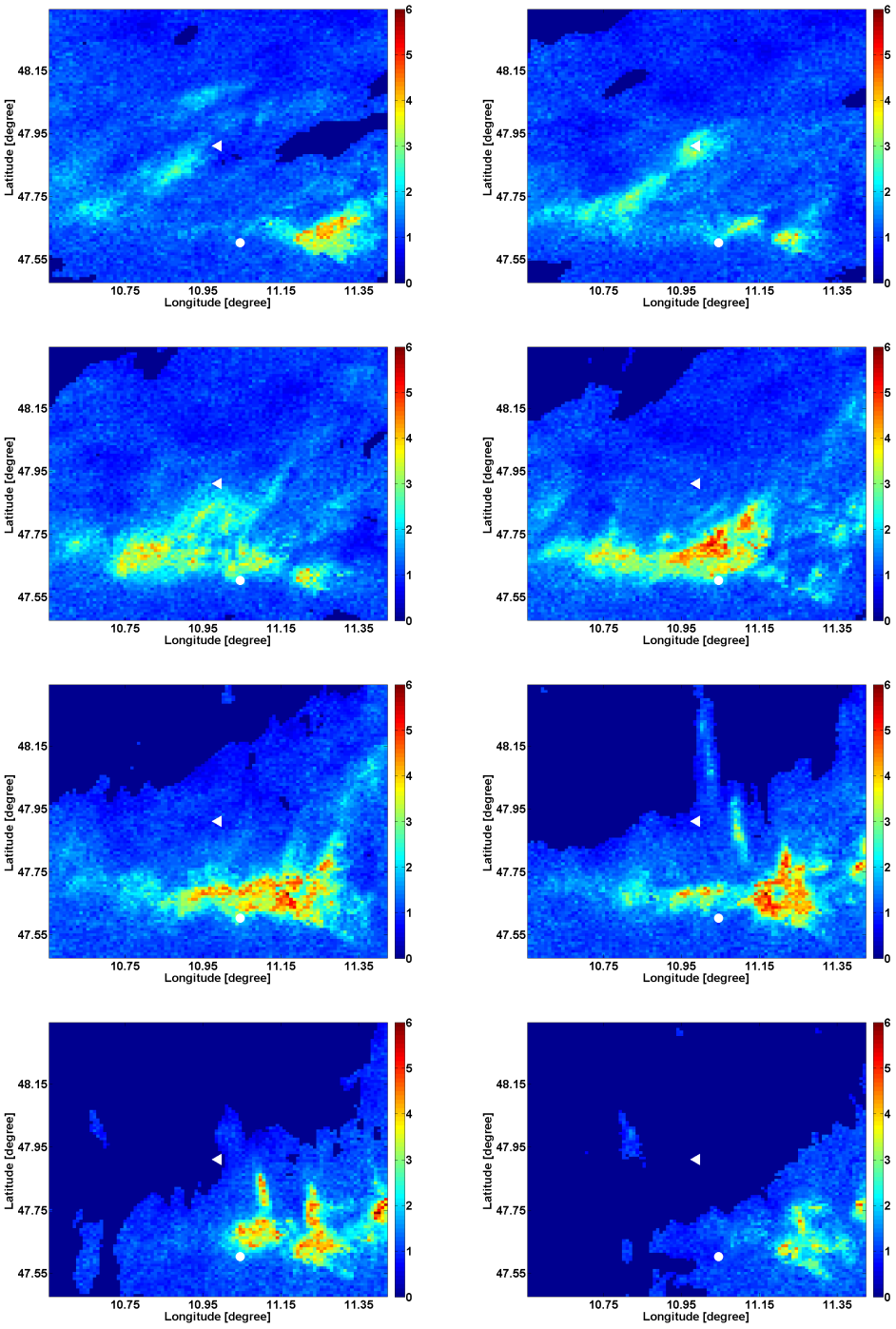

Figure 7.13: Simulated field of pseudo observations derived by the *Maximum Theta* approach for a complete rainfall event from 09:00 to 16:00 (from left to right, top to bottom), 04.08.2008, Gumbel Copula, w/ ARMA-GARCH transformation. The unit of the color bar is in [mm/hour].

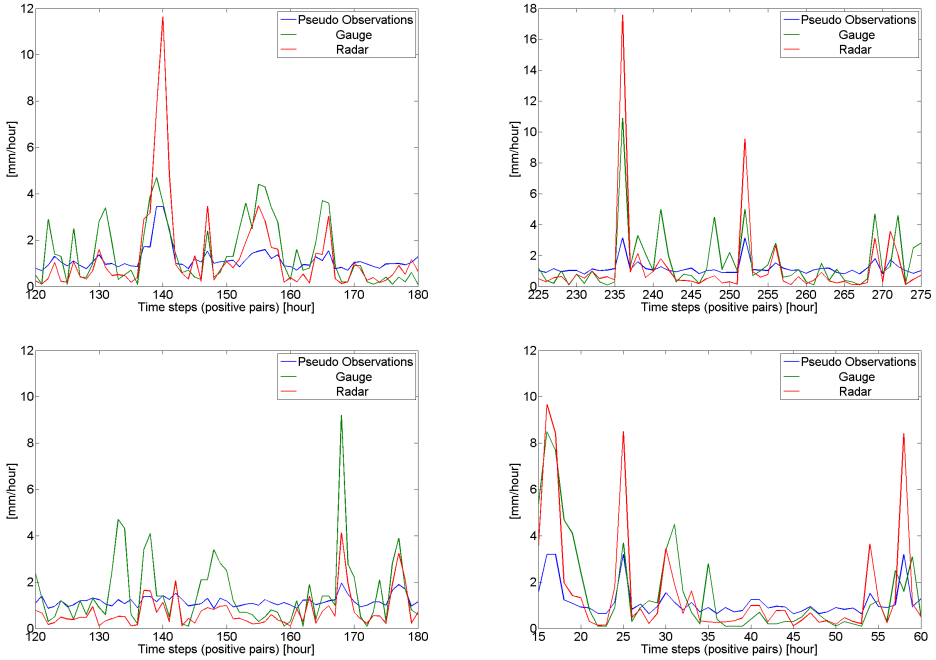

Figure 7.14: Time series of radar (Z/R-256/1.42), gauges and pseudo observations generated by using *Maximum Theta* approach, positive pairs, at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich (from left to right and top to bottom), summer, 2008, Gumbel Copula, w/ ARMA-GARCH transformation.

7.12. Comparing to the original radar field as can be seen in the bottom of Figure 7.4, the basic spatial distribution of rainfall is also remained. However, because of the reduced Copula parameter, the simulated pseudo observation fields are more smoothed than the estimated precipitaiton fields in Figure 7.4 and Figure 7.7 derived from the *Maximum Theta* and *Multiple Theta* approaches.

Additionally, in Figure 7.13, the simulated fields of pseudo observation are presented for the complete rainfall event from 08:00 to 16:00, 04.08.2008, by using the *Maximum Theta* approach.

The example plots of simulated time series of pseudo observation derived from the *Maximum Theta* approach w/ ARMA-GARCH transformation, the corresponding radar and gauge observations are presented in Figure 7.14. Compared to the results listed in Figure 7.10, the simulated time series of pseudo observation are more smoothed, with the compressed peaks. Therefore, the under-estimations in the time se-

Table 7.3: Point wise cross-validation of the *Multiple Theta* and *Maximum Theta* approach for all stations, summer 2008, positive pairs only, Gumbel Copula, w/ ARMA-GARCH transformation.

		Multiple Theta				Maximum Theta			
ID	Station Name	τ	r	RMSE	NSE	τ	r	RMSE	NSE
		[–]	[–]	[mm/h]	[–]	[–]	[–]	[mm/h]	[–]
1	Bernbeuren-Prachtsried	0.58	0.74	1.87	0.07	0.53	0.73	1.80	0.14
2	Diessen	0.49	0.78	1.89	0.08	0.47	0.78	1.80	0.16
3	Deisenhofen	0.35	0.61	2.10	0.03	0.38	0.58	2.04	0.09
4	Ettal	0.35	0.61	1.96	0.04	0.38	0.63	1.93	0.08
5	Garmisch-Partenkirchen	0.52	0.61	2.00	0.09	0.49	0.61	1.90	0.18
6	Gilching	0.42	0.75	2.09	0.08	0.34	0.69	1.20	0.16
7	Griesen	0.45	0.69	1.77	0.10	0.42	0.66	1.71	0.16
8	Halblech	0.48	0.67	2.13	0.07	0.46	0.65	2.03	0.15
9	Hindelang	0.32	0.53	2.30	0.01	0.36	0.58	2.16	0.12
10	Hohenpeissenberg	–	–	–	–	–	–	–	–
11	Kaufbeuren	0.45	0.75	1.51	0.10	0.43	0.73	1.40	0.23
12	Kochel	0.52	0.73	1.81	0.10	0.46	0.68	1.72	0.17
13	Kohlgrub, Bad	0.51	0.77	2.07	0.05	0.46	0.73	1.95	0.16
14	Kraftisried	0.31	0.61	2.13	0.03	0.31	0.57	2.03	0.12
15	Kreuth	0.44	0.66	2.23	0.04	0.42	0.65	2.10	0.15
16	Krün	0.48	0.68	2.30	0.10	0.45	0.64	2.14	0.20
17	Lenggries	0.47	0.69	2.12	0.07	0.46	0.68	1.88	0.27
18	Maisach	0.30	0.70	1.88	0.07	0.30	0.69	1.78	0.16
19	Marktoberdorf	0.58	0.74	2.13	0.06	0.55	0.74	2.03	0.15
20	München	0.42	0.56	1.73	0.08	0.43	0.60	1.65	0.16
21	Oberammergau	0.52	0.73	1.87	0.10	0.46	0.74	1.80	0.17
22	Oberschleissheim	0.27	0.56	2.19	0.03	0.38	0.60	2.12	0.10
23	Oy	0.35	0.50	1.98	0.10	0.39	0.57	1.88	0.10
24	Schwangau	0.43	0.61	1.80	0.07	0.39	0.60	1.71	0.16
25	Seeg	0.53	0.67	2.18	0.05	0.49	0.69	2.03	0.17
26	Schäftlarn	0.53	0.75	1.84	0.05	0.48	0.72	1.75	0.15
27	Steingaden	0.64	0.88	2.33	0.11	0.56	0.87	2.06	0.30
28	Schwaben	0.41	0.45	2.10	0.04	0.37	0.42	2.08	0.05
29	Schlehdorf	0.45	0.64	2.31	0.02	0.46	0.67	2.18	0.13
30	Vilgertshofen	0.52	0.67	2.08	0.09	0.49	0.66	1.97	0.19
31	Wielenbach	0.40	0.63	1.90	0.08	0.36	0.63	1.81	0.16

ries of pseudo observations are enlarged compared to the results w/o ARMA-GARCH transformation, which is also supported by the point wise cross-validation results as listed in Table 7.3. Compared to the validation results presented in Table 7.1 and Table 7.2, the mean values for Kendall's τ is reduced for about 20 percent and reduced around 14 percentfor the Pearson's correlation coefficient. The mean values for RMSE and NSE are also reduced for around 9 percent and 43 percent separately. So, it can be concluded that the point wise cross-validation results are worse by including the ARMA-GARCH transformation.

7.5 Summary and Discussion

The Copula parameter maps presented in this chapter are based on the pairwise Copula analysis. Furthermore, the anisotropies are also incorporated in the dependence structure by combination of the derived Copula parameter maps. As shown in Figure 7.1, the isotropy assumption is not realistic so that the Copula parameter maps with more details of the dependence structure are introduced in this study. At specific gauge stations (e.g. in Garmisch-Partenkrichen, see Fig.7.2, the strong asymmetric structure can be found in the calculated Copula parameter θ_{G} map, indicating the impacts from orography. This asymmetry is disregarded by standard interpolation methods such as inverse distance weighting although it is getting more important in complex terrain (Vogl et al., 2012). Asymmetries in the dependence structure can also be an indicator of air flow directions dominant for a certain location (Vogl et al., 2012). This theory could be supported by investigation of different time scales. For small time scales ($<$ 1 hour), localized rainfall events are not resolved by the statistical analysis. In that case, the region of strongly correlated grid cells, visualized through high Copula parameters in the Copula parameter map, is expected to be reduced compared to larger time scales. For different seasons different types of rainfall regimes are predominant. Therefore, the investigation of e.g. summer and winter season separately is expected to reveal preferential precipitation types, differentiating large scale winter precipitation from convective summer events.

Based on the Copula parameter maps, two Copula-based spatial data assimilation approaches, the *Multiple Theta* and the *Maximum Theta*, are proposed to integrate radar and gauge precipitation information at the spatial scale. Both results, w/o and

w/ ARMA-GARCH transformation, are presented in this chapter. It is found that the simulation results w/ ARMA-GARCH transformation become worse, because of the reduced Copula parameters. Then, the under-estimations in the simulated time series of pseudo observations are enlarged.

The *Maximum Theta* approach is found to be slightly better than the *Multiple Theta* method. This is due to the fact that for each radar grid cell the simulation is based on the highest gauge/radar dependence structure and the respective marginal distribution while for the *Multiple Theta* case stations with low correlation also slightly contribute to the simulated results.

As all ingredients for the proposed algorithms namely the marginal distributions and the theoretical Copula functions only have to be estimated once the proposed bias-correction is computationally not very demanding which facilitates operational application of the proposed methods in quasi real time (Vogl et al., 2012).

Finally it is shown that those proposed Copula based spatial assimilation approaches are able to combine the advantages of the two data sources: the rainfall patterns observed by the radar measurements are retained in the simulated field while the absolute values are successfully corrected towards the gauge observations (Vogl et al., 2012). Then the conclusion can be made that the Copula model is suitable to merge the advantages of the different data sources: the spatial distribution of the radar rainfall field is preserved while absolute values are corrected towards gauge observations (Vogl et al., 2012).

8

Data Assimilation Approach Based On Interpolated Copula Parameter Field

8.1 Introduction

In this chapter, the second Copula-based data assimilation approach is further developed to combine radar and gauge precipitation information, which is also an extension from point wise data assimilation to the spatial case.

As reported e.g. in Bouilloud et al. (2010) and Wagner et al. (2004), radar measurements are dependent on the range to the radar station as well as the local altitude, which is also implied by the classical radar equation liseted below (P_t is the transmitter power; G_t is the gain of the transmitting antenna; A_r is the effective aperture of the receiving antenna; σ is the radar cross section, or scattering coefficient, of the target; F is the pattern propagation factor; Ra is the range to the radar)

$$P_r = \frac{P_t G_t A_r \sigma F^4}{(4\pi)^2 Ra^4} \tag{8.1}$$

Far away from the radar station, the uncertainties in the radar observations are larger as the radar beams become larger. Furthermore, the the received power P_r is also smaller due to the increase of Ra. Additionally, with the increasing altitude, the temperature decreases. This can lead to the change of precipitation type, e.g. from rain to snow at different altitudes. Bourgouin et al. (2000) also found that the temperature can have an impact on the precipitation type.

With the assumption that the radar/gauge dependence structure - Copula parameter θ_G - is dependent on the range to the radar station and the local altitude for each grid cell, the Copula parameter θ_G field can be interpolated for the whole domain. By combining both statistical and physical backgrounds of precipitation process, a virtual rain gauge can be assigned to each grid cell, and finally, the bias corrected spatio-temporal precipitation fields are simulated conditioned on the radar fields. Performing this new Copula-based data assimilation approach also for the Bavarian Alps and alpine forelands, the simulated precipitation fields are compared with the results obtained from the available gauge measurements by means of several different performance indices.

The structure of this chapter is as following. First, in section 8.2, the methodology, another approach to model the spatial distribution of Copula parameter θ_G, is introduced in detail. Then, the interpolated Copula parameter θ_G fields and simulated precipitation fields are presented in section 8.3, for both cases w/o and w/ ARMA-GARCH transformation. Afterwards, the impacts from the interaction between range and altitude to the interpolated Copula parameter θ_G field are discussed in section 8.4. Finally, in section 8.4, the validation procedures are performed.

8.2 Methodology

From theoretical point of view, as already described e.g. in Bouilloud et al. (2010) and Wagner et al. (2004), the radar precipitation measurements are directly influenced by the range to the radar station and local altitude. Meanwhile, the gauge observed precipitation is also affected by altitude. Because the precipitation type may be highly variant in mountain area at the same time step, i.e. rain to snow with the increasing of altitude. Then the assumption can be made that the radar/gauge dependence structure is a function of range and altitude. In this chapter, the polynomial with 2 variables listed below (Ra refers to range to the radar station and z refers to altitude for each grid cell) is used to model the spatial distribution of the dependence structure between radar and gauge.

$$\theta_G = f(Ra, z) = f(Ra) + f(z) \tag{8.2}$$

The order of the variates in the polynomial is determined according to the residuals calculated between the fitted and real Copula parameter θ_G values, which is the

reference to test the quality of fit. Note that, with the increasing order of the variates in the polynomial, the unreasonable extreme large or small values of $f(Ra, z)$ can be found. This is also the main disadvantage for using polynomial here, which can lead to the totally wrong representations for the spatial distributions of the dependence structure and has to be considered carefully as discussed in 8.3, even though the residuals for those cases are small. As a summary, the references to determine the order of the variates in the polynomial are listed below:

1. The residuals calculated between the fitted and real Copula parameter θ_{G} values.
2. The final Copula parameters calculated according to $f(Ra, z)$ should be not be extreme large or small especially for the cases with small values for Ra and large values for z.

Provided with the interpolated Copula parameter θ_{G} field, the dependence structure between radar and gauge can be identified for each grid in the whole domain. Then the pseudo observation field can be simulated conditioned on the given radar measurement for each time step. The basic steps for simulation of pseudo observations based on the interpolated Copula parameter θ_{G} field can be summarized as following:

1. Based on the interpolated Copula parameter θ_{G} field, the radar/gauge dependence structure is identified for each specific radar grid cell.
2. One sample of 100 members is simulated in the rank space conditioned on the given radar measurement.
3. The mean value is calculated from the random samples.
4. The rain gauge marginal distribution can be selected the same as in the *Maximum Theta* approach.
5. The integral transformation is applied on the mean value to transform back to the data space.
6. These steps are repeated for all radar grid cells to generate the precipitation field of pseudo observation.

8.3 Results

By performing the method introduced in the last section, the interpolated Copula parameter θ_G field can be calculated and then the pseudo observation fields can be further simulated. The simulation results are presented in this section, for both cases w/o and w/ ARMA-GARCH transformation.

8.3.1 Interpolated Copula Parameter Field

The foundation for this new Copula-based spatial data assimilation approach is the interpolated Copula parameter θ_G field. Once the Ra and z are known, the Copula parameter θ_G is interpolated for each grid cell in the domain, showing the magnitude of dependence between gauges and the corresponding radar measurements.

8.3.1.1 Without ARMA-GARCH Transformation

As already introduced in section 8.2, the polynomial with 2 variables - Ra and z - is used to represent the range and altitude dependent relationship. However, the difficulty is to find the suitable order for the variates in the polynomial by using the residuals between the fitted and real Copula parameter θ_G values, as well as the principle to avoid the unreasonable values for $f(Ra, z)$.

In order to test the quality of fitting results, the Pearson's correlation and RMSE are calculated between the fitted and real Copula parameter θ_G values. Note that the RMSE is used here only to test the match between the fitted and real Copula parameter values, and therefore has no real physical meaning and no also unit. As shown in Figure 8.1, with the increasing of the orders for both Ra and z, the fitted results becomes better as the Pearson's correlation r increasing from 0.6 to near 1 and RMSE decreasing from 0.09 to 0.05.

However, when the order of Ra is larger than 4 or the order of z is larger than 2 (the order for the other variate is fixed), unreasonable extreme large or small values are calculated especially at the with small Ra or with large z as shown in Figure 8.2. Therefore, the orders for Ra and z are chose at 4 and 2 respectively. As a result, the $f(Ra, z)$ listed as below is used to calculate the Copula parameter θ_G based on Ra and z for each grid cell in the case w/o ARMA-GARCH transformation.

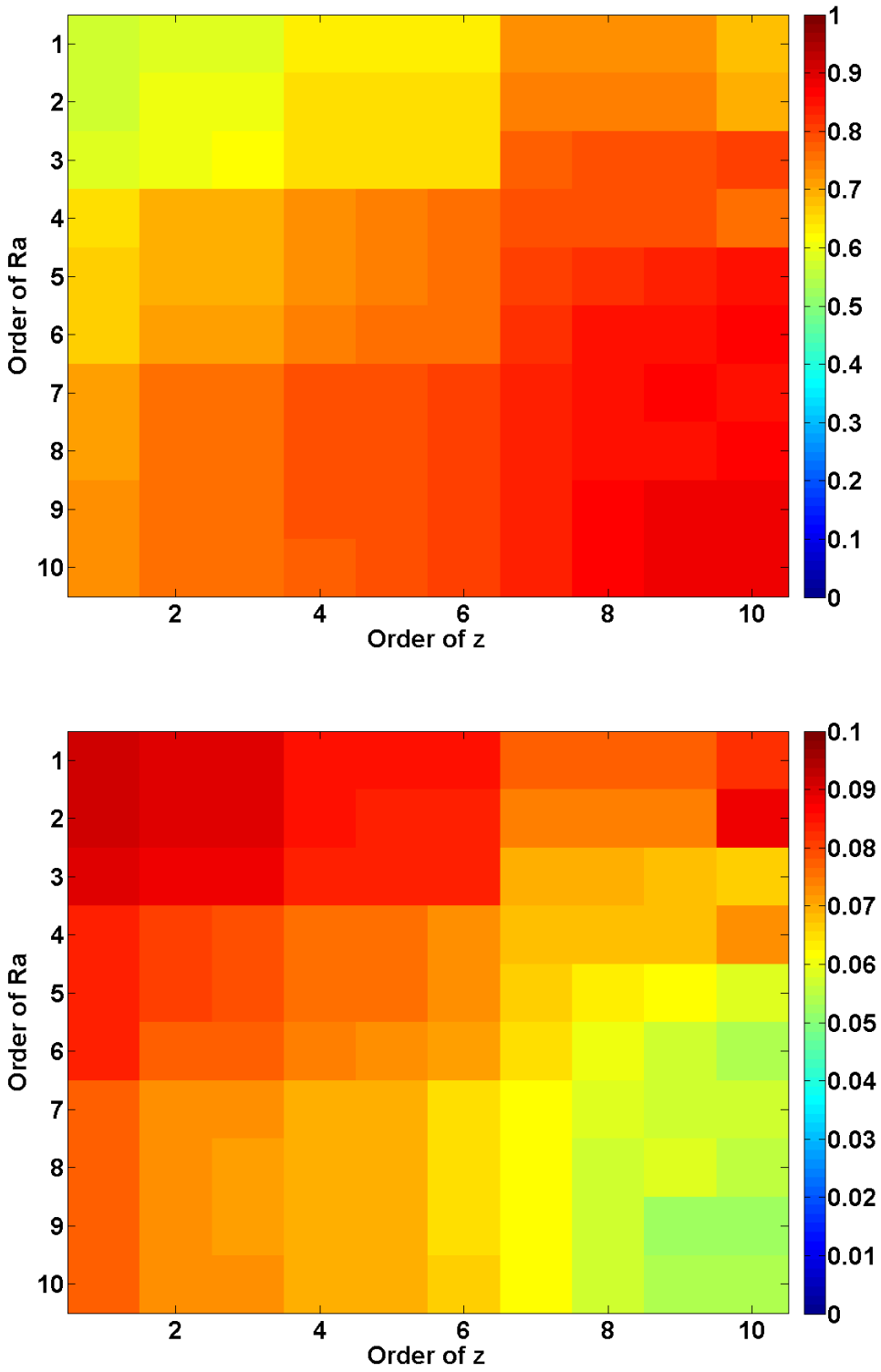

Figure 8.1: The Pearson's correlation r (top) and RMSE (bottom) calculated between the fitted and real Copula parameter θ_G values, by using the polynomials under different orders for Ra and z, summer, 2006 to 2007, w/o ARMA-GARCH transformation.

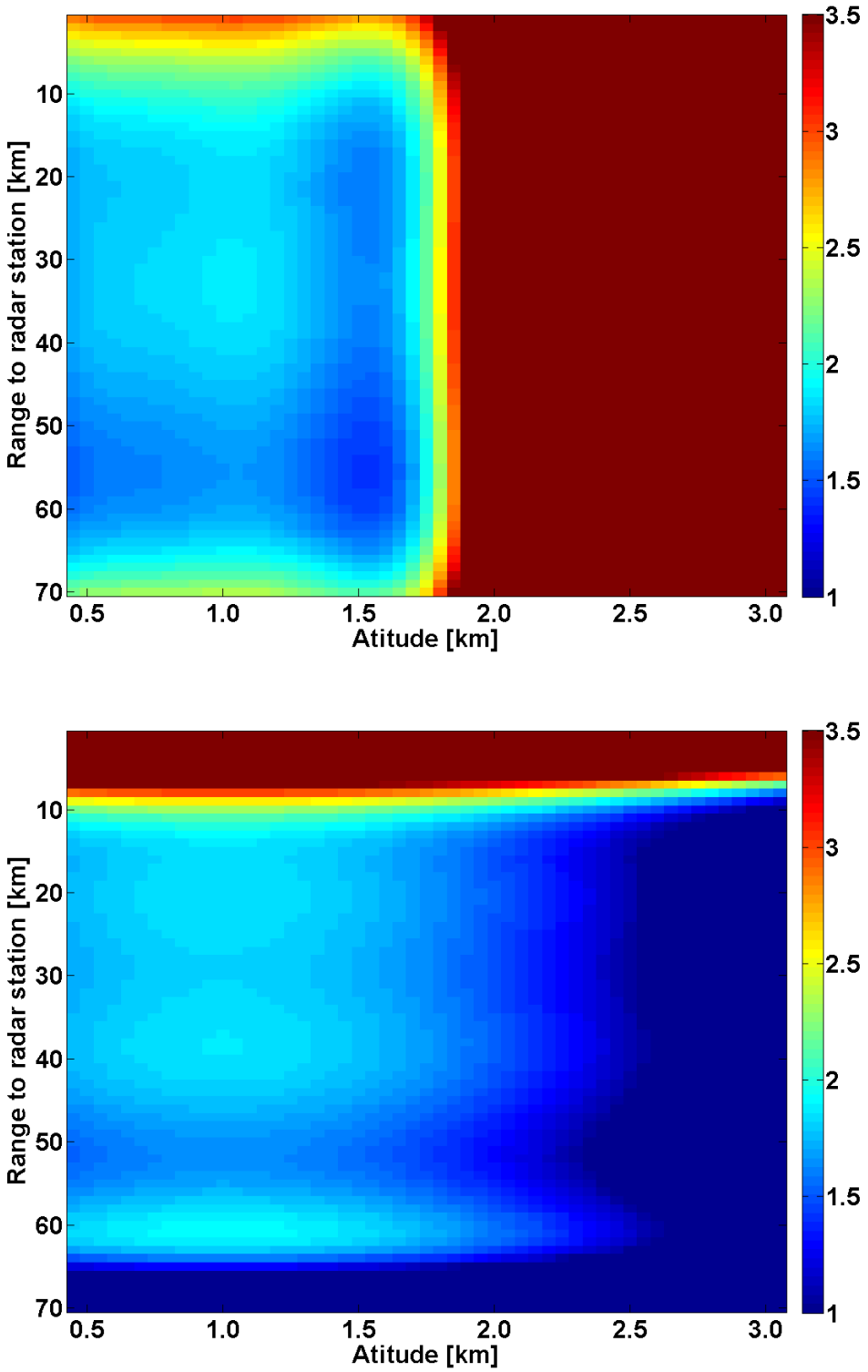

Figure 8.2: Copula parameter θ_G calculated according to Ra and z (the orders for Ra at 4 and z at 5 in the top; the orders for Ra at 7 and z at 2 at the bottom), summer, 2006 to 2007, w/o ARMA-GARCH transformation.

$$\theta_{\mathrm{G}} = f(Ra, z) = c_0 + \sum_{i=1}^{4} c_i Ra^i + c_5 z + c_6 z^2 \tag{8.3}$$

Table 8.1: Coefficients of the fitted two variates polynomial, summer, 2006 to 2007, w/o ARMA-GARCH transformation.

Coefficient	c_0	c_1	c_2	c_3
Value	2.79	-0.17	0.008	-0.00016
Coefficient	c_4	c_5	c_6	-
Value	-0.0000011	0.58	-0.32	-

The fitted coefficient values for $f(Ra, z)$ are listed in Table 8.1 and the corresponding residuals calculated between the fitted and real Copula parameter θ_{G} values are shown in the Table 8.2. The maximum one (absolute value) is -0.20 and the second largest is 0.15 as highlighted, about 8 to 10 percent of the corresponding real θ_{G}. Meanwhile, the minimum one (absolute value) is 0.01, less than 0.2 percent of the real θ_{G}. The majorities (27 in total 30 which means 90 percent) are around 3 to 5 percent of the real θ_{G}. So, it is concluded that this polynomial can represent the spatial distribution of radar/gauge dependence structure reasonably well.

Table 8.2: Residuals between between the fitted and real Copula θ_{G} values (arranged by the number in Table 4.1, ID number 10 of Hohenpeissenberg is not included), summer, 2006 to 2007, w/o ARMA-GARCH transformation.

Gauge ID	1	2	3	4	5	6	7	8	9	11
Residual	-0.10	0.04	-0.03	-0.11	-0.14	0.09	0.01	0.07	-0.04	0.03
Gauge ID	12	13	14	15	16	17	18	19	20	21
Residual	-0.05	-0.04	0.02	0.11	**-0.20**	0.05	0.07	0.08	-0.12	0.03
Gauge ID	22	23	24	25	26	27	28	29	30	31
Residual	0.02	0.05	0.11	**0.15**	-0.06	-0.02	-0.02	-0.03	0.01	0.04

The Copula parameter θ_{G} values calculated for different range and altitude are shown in the top of Figure 8.3 (summer, 2006 to 2007). It can be seen that, in general,

the Copula parameter θ_G is decreased with the increasing range and increasing altitude. This matches with the fact that the uncertainty in the radar should be larger with the increasing Ra and z. However, when the range is larger than 65 km, θ_G is increasing with the increasing Ra. This can be due to the properties of the polynomials used here and also due to the fact that no real gauges located with the range larger than 65 km.

Similarly, since there are also limited number of gauges or even no gauges in the region with high altitudes, the calculated Copula parameter θ_G is even smaller than one (the possible smallest value for θ_G) if $z > 2500$ m.a.s.l. To overcome this problem, the final Copula parameter θ_G is set to be 1 if the fitted one is smaller than 1. In this study area, there are only 14 grid cells with the altitude larger than 2500 m.a.s.l (0.14 percent of the total 10000 grid cells) so that the influence to the final results is also limited as shown in the middle of of Figure 8.3. However, one disadvantage of this approach is dependent on the distribution of the rain gauges along Ra and z, but not the total rain gauge number.

Once the range to the radar station and local altitude are known, Copula θ_G can be calculated for each grid cell in the whole domain as shown in the middle of Figure 8.3, together with digital elevation model (DEM) at the bottom. It can be seen that close to the radar station, θ_G is large and mainly determined by the range. This is not only due to the relatively low absolute altitudes but also due to the small altitude gradient at those flat areas. By contrast, in the southern part - Alps and alpine forelands, both the altitude and its gradient become much larger as can be seen from the DEM model. Therefore, the distribution of Copula parameter θ_G is mainly dependent on the altitude and shows a pattern similar as in the DEM model.

Additionally, the interpolated range and altitude dependent Copula parameter θ_G fields for spring, autumn and winter are presented in Figure 8.4. Different temperature can lead to variant precipitation types so that those interpolated Copula parameter θ_G fields are different from each other and show different patterns.

8.3.1.2 With ARMA-GARCH Transformation

Similar as in the approach w/ ARMA-GARCH transformation, with the increasing of the orders for both Ra and z, the fitted results becomes better as the Pearson's correlation r increasing from 0.4 to 0.8 and RMSE decreasing from 0.09 to 0.05, as shown in Figure 8.5.

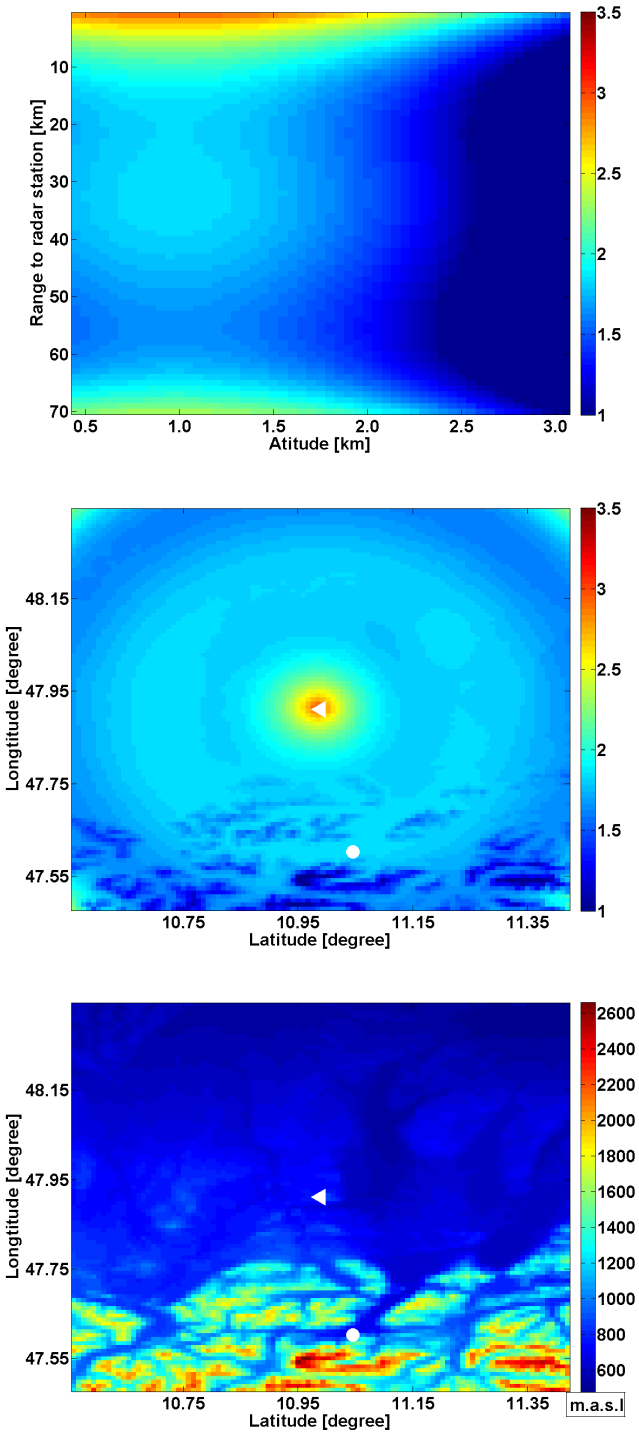

Figure 8.3: Copula parameter θ_{G} as a function of range and altitude (top), interpolated Copula parameter θ_{G} field (middle) in summer 2006 to 2007 and digital elevation model (bottom), w/o ARMA-GARCH transformation. The color bar is the value of the Copula parameter θ_G. The white triangle refers to the radar station at Hohenpeissenberg and the white cycle refers to Garmisch-Partenkirchen, the same all through the thesis.

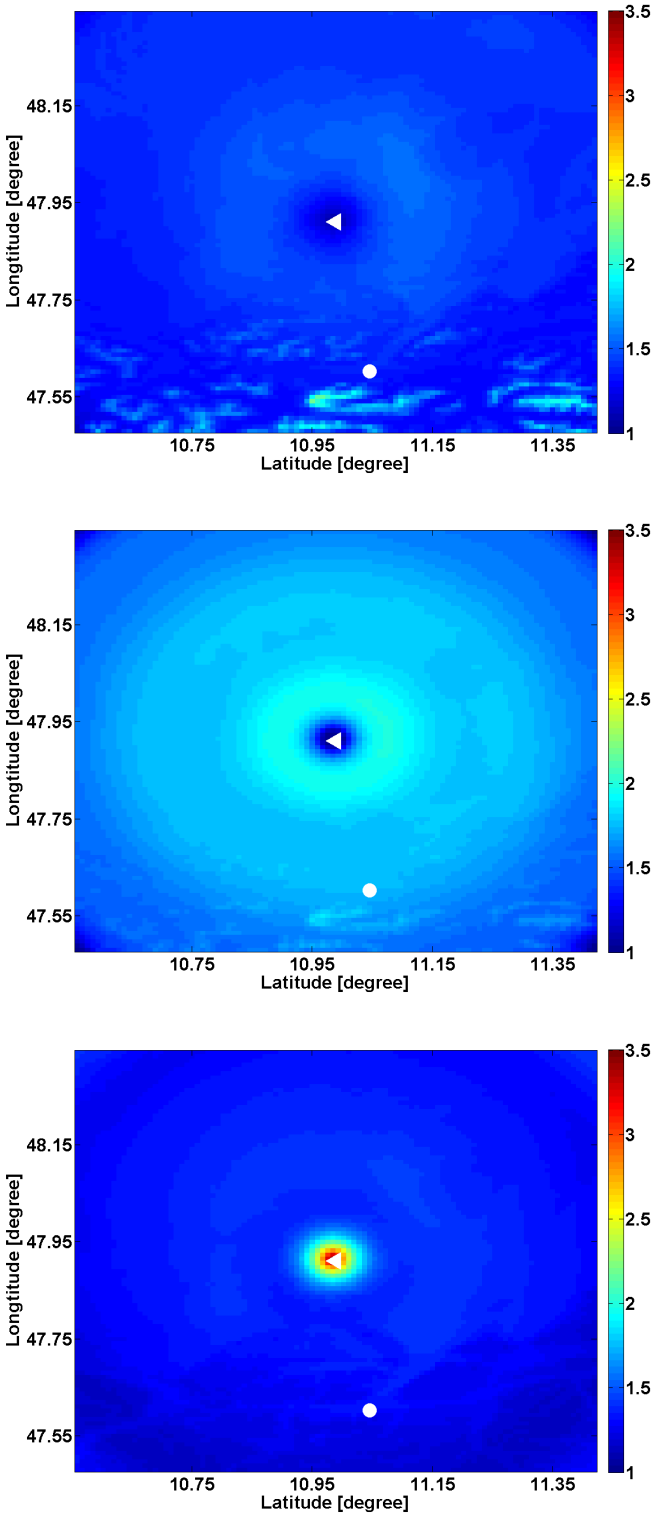

Figure 8.4: Interpolated Copula parameter θ_{G} fields for spring, autumn and winter (top to bottom), 2006 to 2007, w/o ARMA-GARCH transformation. The color bar is the value of the Copula parameter θ_G.

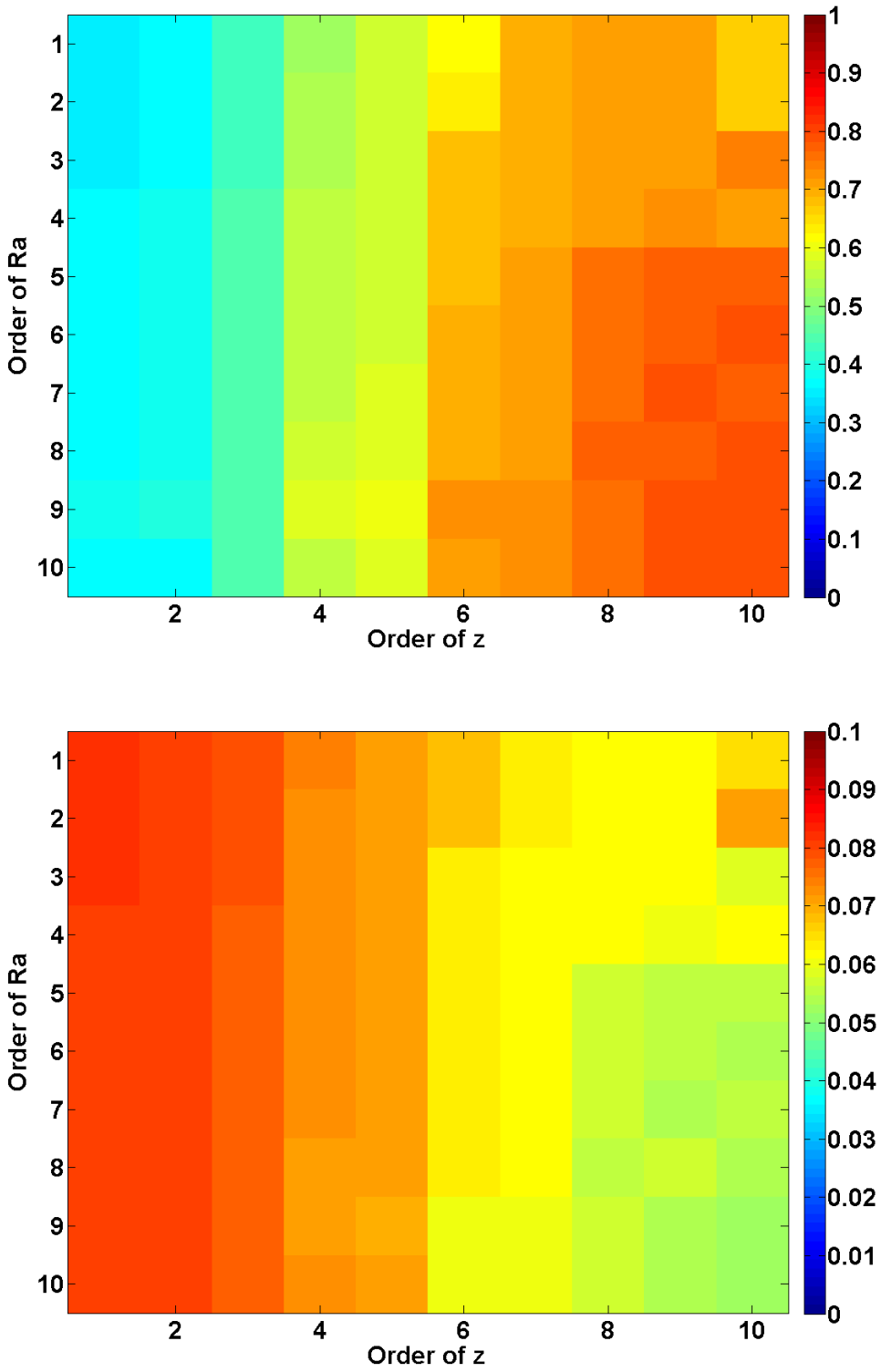

Figure 8.5: The Pearson's correlation (top) and rmse (bottom) calculated between the fitted and real Copula parameter θ_G values, by using the polynomials under different orders for Ra and z, summer, 2006 to 2007, w/ ARMA-GARCH transformation.

However, when the order of Ra is larger than 3 or the order of z is larger than 1 (the order for the other variate is fixed), unreasonable extreme large or small values are calculated especially at the locations close the radar station or with high altitude similar as shown in Figure 8.2. Therefore, the orders for Ra and z are chose at 3 and 1 respectively and the $f(Ra, z)$ is listed as below:

$$\theta_{\mathrm{G}} = f(Ra, z) = c_0 + \sum_{i=1}^{3} c_i Ra^i + c_4 z \tag{8.4}$$

Table 8.3: Coefficients of the fitted bivariate polynomial, summer, 2006 to 2007, w/ ARMA-GARCH transformation.

Coefficient	c_0	c_1	c_2	c_3	c_4
Value	1.42	-0.002	0.000021	-0.00000038	0.04

The fitted coefficient values for $f(Ra, z)$ are listed in Table 8.3 and the corresponding residuals between the fitted and real θ_{G} value is also shown in the Table 8.4. The maximum residual (absolute value) is -0.18 and the second largest is -0.16 as highlighted, about 11 percent of the corresponding real θ_{G}. Meanwhile the minimum (absolute value) is 0.01, less than 0.5 percent of the real θ_{G}. The majorities (27 in total 30 means 90 percent) are around $3 \sim 10$ percent of the real θ_{G}. So, it is also concluded that this polynomial can represent the spatial distribution of radar/gauge dependence structure when the ARMA-GARCH transformation is included.

The theoretical range and altitude dependent Copula θ_{G} function is shown in the top of Figure 8.6 for summer 2006 to 2007, as well as the interpolated Copula parameter θ_{G} field at the bottom. Similar as the results in 8.3.1, it can be seen that, in general, θ_{G} is decreased with the increasing range and also with the increasing altitude. The main difference is still the relatively smaller Copula parameter value compared to the results w/o the ARMA-GARCH transformation. The problems discussed in 8.3.1 are still remained here.

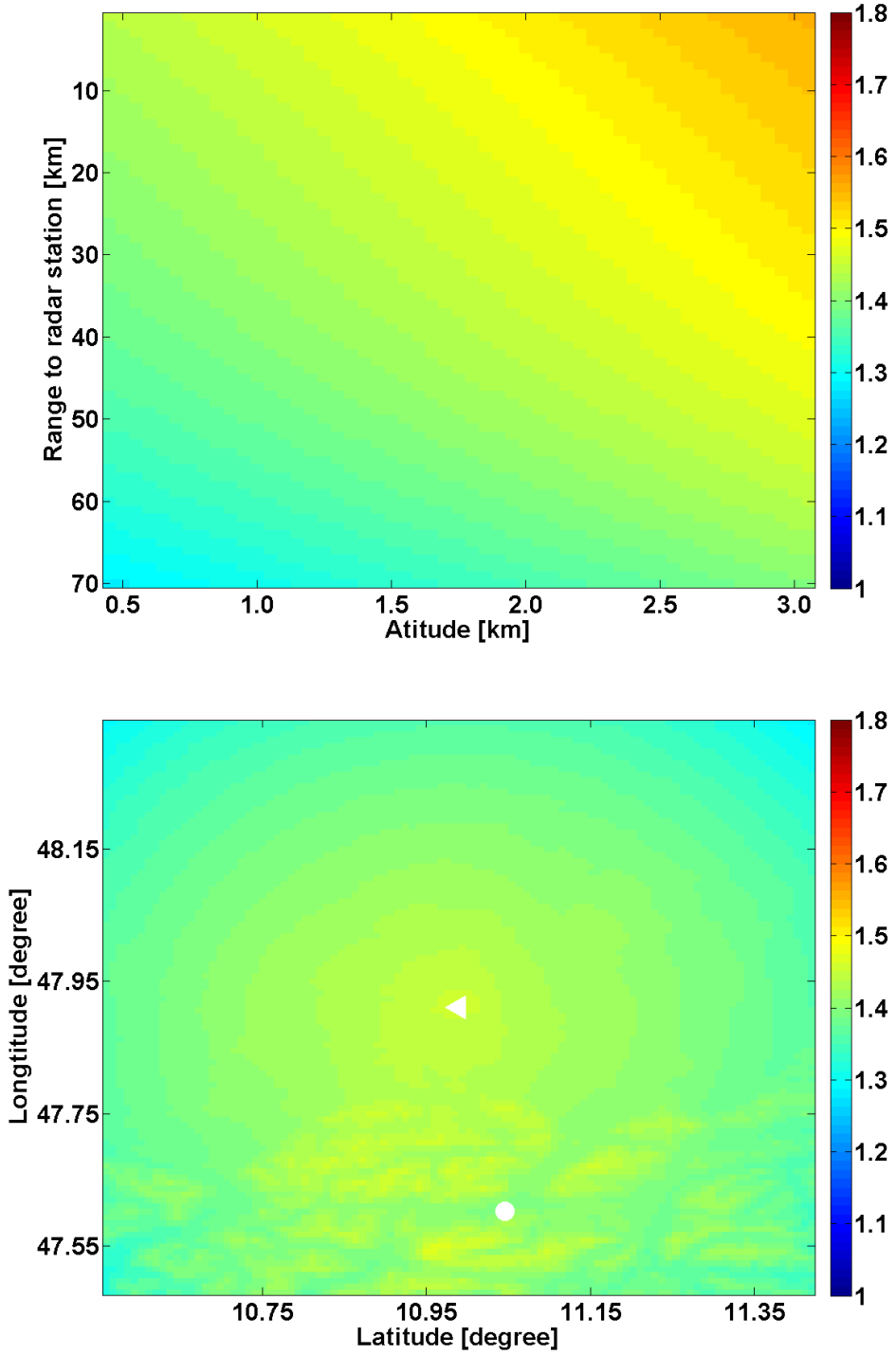

Figure 8.6: Copula parameter θ_G as a function of range and altitude (top) and the interpolated Copula parameter θ_G field (bottom), summer, 2006 to 2007, w/ ARMA-GARCH transformation. The color bar is the value of the Copula parameter θ_G.

Table 8.4: Residuals between between the fitted and real Copula θ_G value (arranged by the number in Table 4.1, ID number 10 of Hohenpeissenberg is not included), w/ ARMA-GARCH transformation.

Gauge ID	1	2	3	4	5	6	7	8	9	11
Residual	0.12	0.05	-0.08	-0.07	0.01	0.02	0.04	0.08	-0.02	0.02
Gauge ID	12	13	14	15	16	17	18	19	20	21
Residual	-0.11	-0.07	0.07	0.04	**-0.18**	0.01	0.10	0.03	-0.01	-0.03
Gauge ID	22	23	24	25	26	27	28	29	30	31
Residual	0.01	0.05	0.15	0.06	0.05	0.16	-0.12	-0.13	**-0.16**	0.04

8.3.2 Simulated Field of Pseudo Observations

An example of precipitation field generating by applying the interpolated Copula parameter θ_G field based approach is shown in the top of Figure 8.7. Comparing this simulated field of pseudo observations with the uncorrected radar field (see at the bottom of Figure 7.4), it can be seen that the basic patten from the original radar field are preserved in a reasonable way and the absolute values are corrected towards the ground measurements. Generally, this simulated field is close to the precipitation field derived from *Maximum Theta* approach (see Figure 7.7). However, some small differences can still be found e.g. the region with the latitude around 47.75 and longitude around 11.30.

Similarly, the simulated field of pseudo observations w/ ARMA-GARCH transformation can also be generated as presented in the bottom of Figure 8.7. The basic spatial rainfall pattern can also be remained. However, comparing to the simulation result in the top of Figure 8.7, the values in the newly simulated rainfall field are smaller also because of the smaller Copula θ_G values in the interpolated θ_G field after performing ARMA-GARCH transformation.

Additionally, in Figure 8.8 and Figure 8.9, the simulated fields of pseudo observation are also presented for the complete rainfall event from 08:00 to 16:00, 04.08.2008, for both cases w/o and w/ ARMA-GARCH transformation.

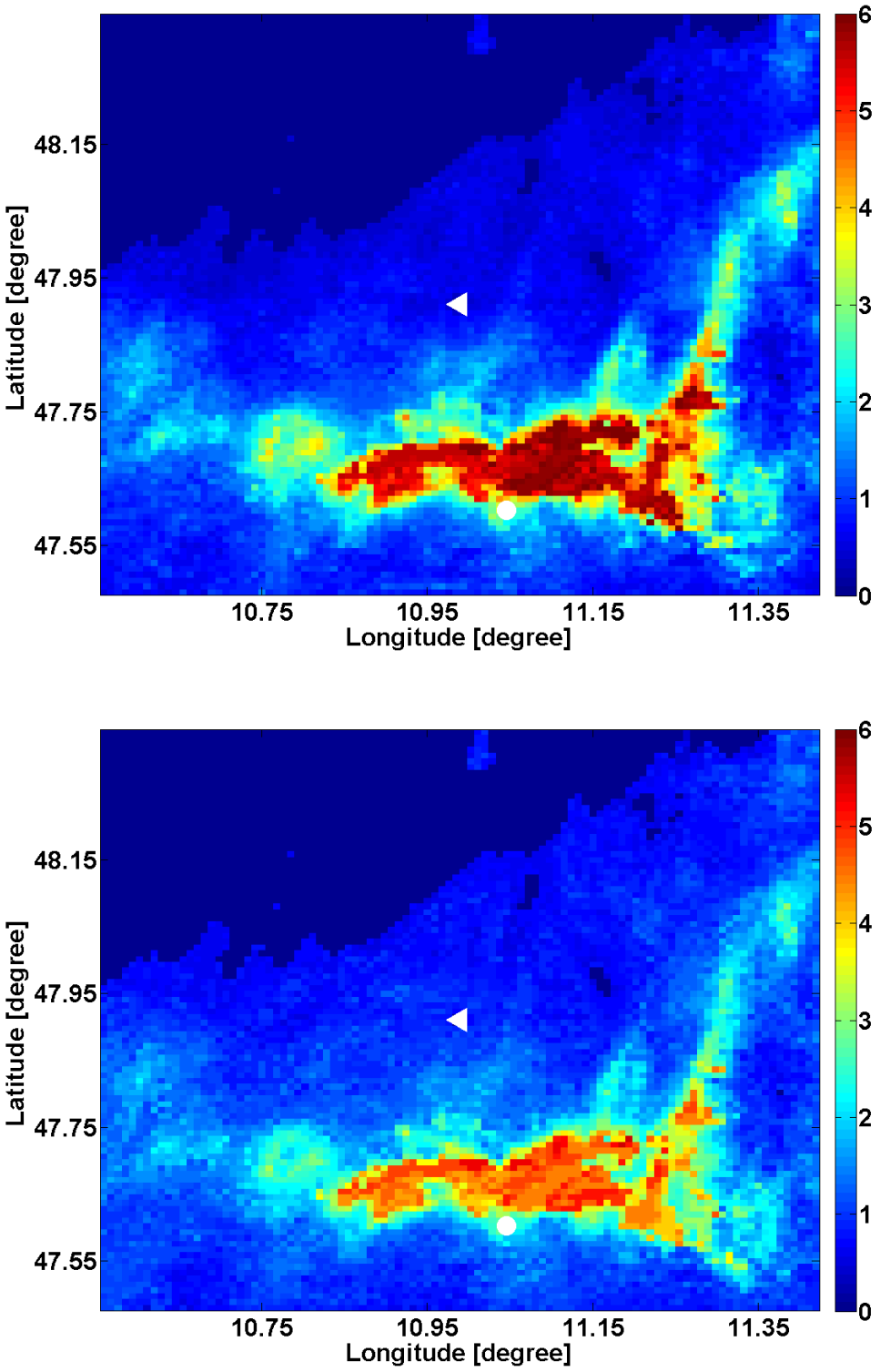

Figure 8.7: Simulated fields of pseudo observations by using the interpolated Copula parameter θ_G field based approach, 13:00, 04.08.2008, w/o and w/ ARMA-GARCH transformation (top to bottom). The unit of the color bar is in [mm/hour].

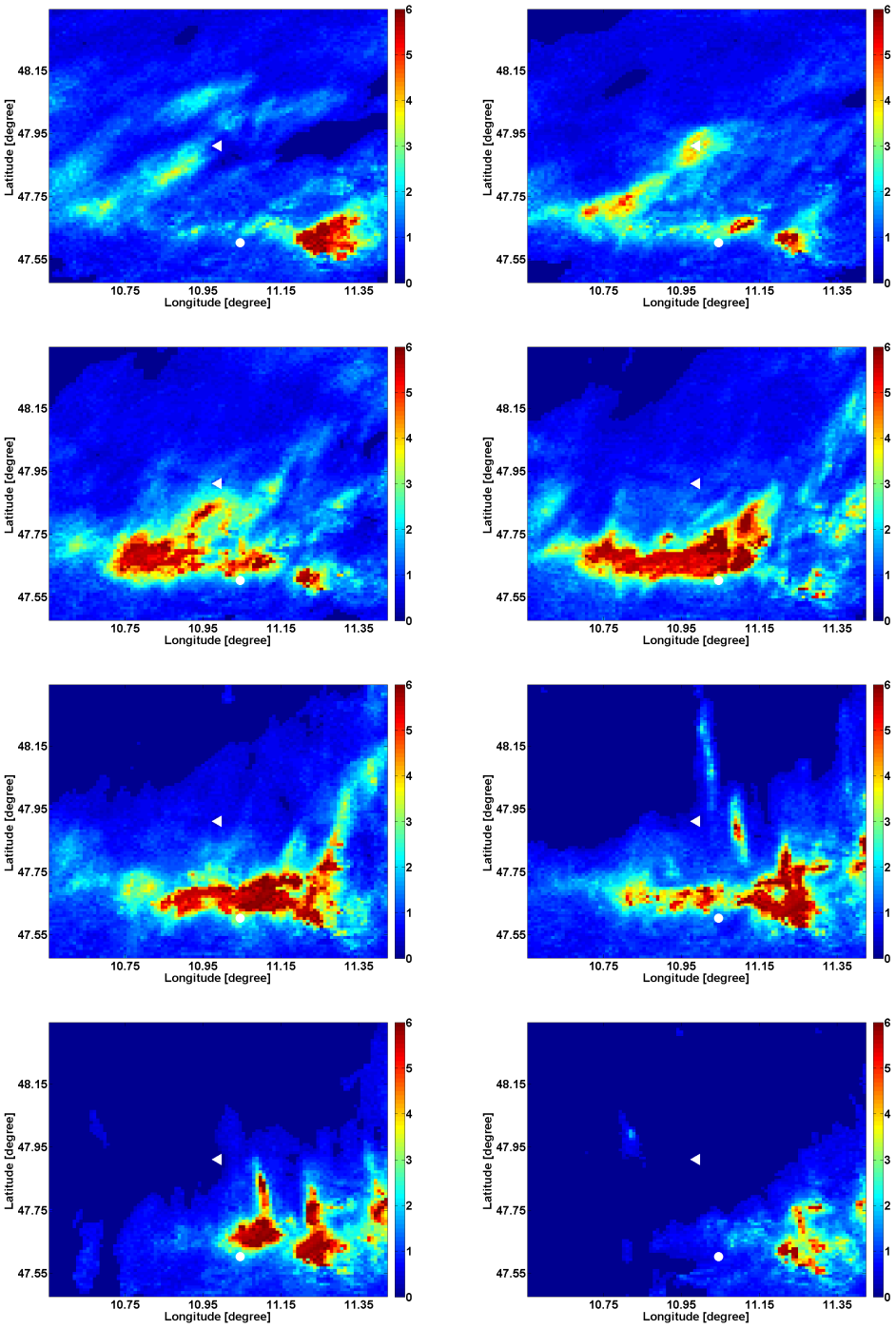

Figure 8.8: Simulated fields of pseudo observations by using the interpolated Copula parameter θ_G field based approach for a complete rainfall event from 09:00 to 16:00 (from left to right, top to bottom), 04.08.2008, w/o ARMA-GARCH transformation. The unit of the color bar is in [mm/hour].

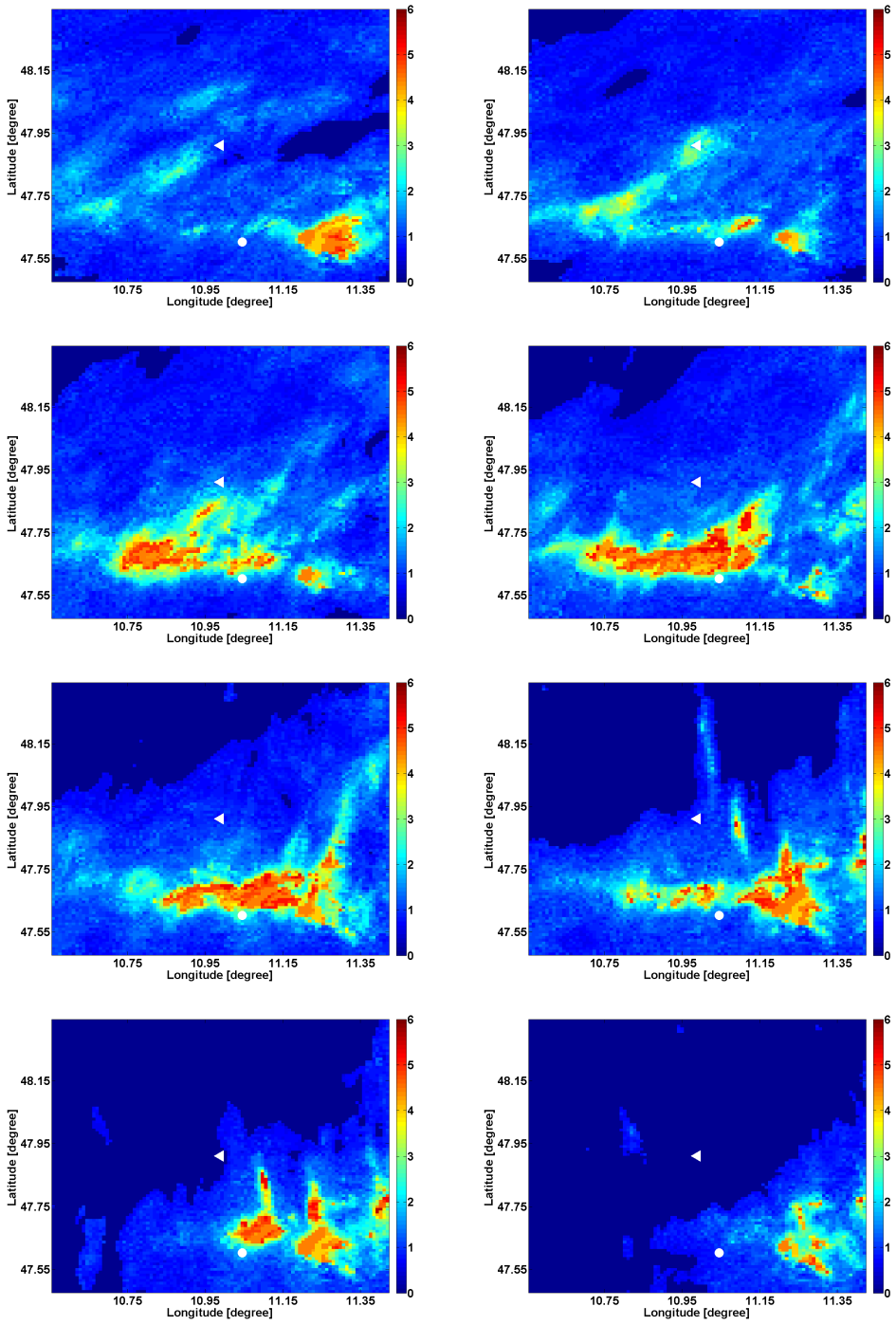

Figure 8.9: Simulated fields of pseudo observation the interpolated Copula parameter θ_G field based approach for a complete rainfall event from 09:00 to 16:00 (from left to right, top to bottom), 04.08.2008, w/ ARMA-GARCH transformation. The unit of the color bar is in [mm/hour].

8.4 Validation of the Simulated Precipitation Fields

In order to check the quality of simulated pseudo observation field, validations are performed and the results are presented in this sub-section, both for the results w/o and w/ ARMA-GARCH.

8.4.1 Without ARMA-GARCH Transformation

Additionally, the simulated time series of pseudo observations, as well as the corresponding gauge and radar observations, are presented in Figure 8.10 at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich during summer 2008, w/o ARMA-GARCH transformation. Similarly as the results shown in Figure 7.10, the over-estimations (the peaks) in the radar observed rainfall are corrected and the under-estimations are still remained for certain time steps, as the simulations are all based on the given radar measurements.

The simulated time series of pseudo observations, by using interpolated Copula parameter θ_{G} field based approach, are analysed by point wise cross-validation using the gauge observations in the respective grid cell and the results are shown in Table 8.5). The mean value for Pearson's correlation between the simulated pseudo observations and the observed precipitation is 0.71, similar as the results by using *Multiple Theta* and *Maximum Theta*, indicating equally good results. However, the results for mean values of RMSE and MAE show that the interpolated Copula parameter θ_{G} field based approach is better, with the improvements at 2 and 4 percent respectively. This finding is also supported by the inspection of the respective NSE value at 0.31, while being 0.28 for the *Maximum Theta* method, nearly 10 percent better. Note that the validation results for the location at Hohenpeissenberg are not included to the mean values as this station is located close to the radar observatory itself.

8.4.2 With ARMA-GARCH Transformation

The point wise cross-validation results for the pseudo observations (generated based on the interpolated Copula parameter θ_{G} field w/ ARMA-GARCH) are listed in Table 8.6, as well as the example simulated time series in Figure 8.11. Compared to the results listed in Figure 8.10, the simulated time series of pseudo observation are more smoothed, with the compressed peaks, leading to the larger under-estimations compared

Table 8.5: Point wise cross-validation for interpolated Copula parameter θ_G field based approach for all the stations, summer, 2008, positive pairs only, w/o ARMA-GARCH transformation.

ID	Station Name	Kendall's τ [–]	r [–]	RMSE [mm/h]	MAE [mm/h]]	NSE [–]
1	Bernbeuren-Prachtsried	0.60	0.73	1.69	0.96	0.24
2	Diessen	0.60	0.81	1.66	0.86	0.29
3	Deisenhofen	0.56	0.66	1.82	0.96	0.27
4	Ettal	0.55	0.68	1.56	0.93	0.39
5	Garmisch-Partenkirchen	0.53	0.63	1.72	0.91	0.32
6	Gilching	0.54	0.79	1.70	0.85	0.39
7	Griesen	0.53	0.70	1.39	0.80	0.44
8	Halblech	0.54	0.69	1.89	0.95	0.26
9	Hindelang	0.50	0.60	2.02	1.10	0.23
10	Hohenpeissenberg	–	–	–	–	–
11	Kaufbeuren	0.62	0.78	1.16	0.72	0.47
12	Kochel	0.57	0.73	1.60	0.91	0.29
13	Kohlgrub, Bad	0.60	0.79	1.72	0.94	0.34
14	Kraftisried	0.58	0.71	1.87	1.00	0.25
15	Kreuth	0.50	0.67	1.98	1.13	0.24
16	Krün	0.54	0.71	1.82	0.96	0.43
17	Lenggries	0.53	0.68	1.73	0.96	0.38
18	Maisach	0.48	0.74	1.55	0.86	0.36
19	Marktoberdorf	0.61	0.75	1.81	1.04	0.32
20	München	0.51	0.62	1.58	0.94	0.23
21	Oberammergau	0.57	0.76	1.60	0.91	0.34
22	Oberschleissheim	0.54	0.77	1.78	0.85	0.36
23	Oy	0.50	0.58	1.69	1.01	0.27
24	Schwangau	0.55	0.64	1.54	0.86	0.33
25	Seeg	0.62	0.68	2.00	1.02	0.20
26	Schäftlarn	0.59	0.79	1.65	0.88	0.24
27	Steingaden	0.67	0.89	1.80	0.89	0.47
28	Schwaben	0.52	0.50	1.90	0.97	0.21
29	Schlehdorf	0.60	0.74	1.98	1.01	0.28
30	Vilgertshofen	0.56	0.70	1.83	0.95	0.30
31	Wielenbach	0.56	0.71	1.63	0.85	0.32

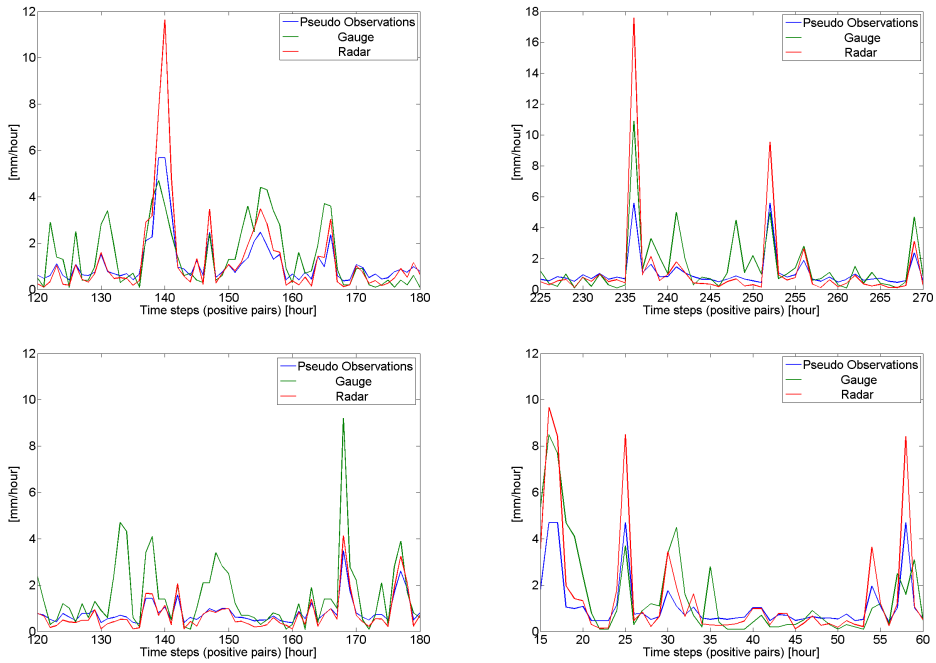

Figure 8.10: Time series of pseudo observations generated by using the interpolated Copula parameter θ_G field based approach, the corresponding radar (Z/R-256/1.42) and gauge observations, for the station Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich (from left to right and top to bottom), positive pairs, summer, 2008, w/o ARMA-GARCH transformation.

to the results w/o ARMA-GARCH transformation. Therefore, similar to the point wise cross-validation results in Chapter 6 and 7, the simulations w/ ARMA-GARCH are worse than those w/o ARMA-GARCH. There are only small differences (reduced from 0.71 to 0.69) on the sense Pearson's correlation coefficient (mean values), the same for Kendall's τ. However, on the sense of RMSE, MAE and NSE (mean values), the results are around 10 to 20 percent worse w/ ARMA-GARCH transformation.

8.5 Summary and Discussion

In this chapter, based on the range and altitude interpolated Copula parameter θ_{G} field, the second Copula-based method is developed to assimilate precipitation information from radar and gauge. Also starting from the pairwise Copula analysis, through integrating the impacts from altitude and range to the radar station, the gauge/radar

Table 8.6: Point wise cross-validation for interpolated Copula parameter θ_G field based approach for all stations, summer, 2008, positive pairs only, w/ ARMA-GARCH transformation.

ID	Station Name	Kendall's τ [–]	r [–]	RMSE [mm/h]	MAE [mm/h]	NSE [–]
1	Bernbeuren-Prachtsried	0.51	0.71	1.77	1.02	0.17
2	Diessen	0.51	0.74	1.75	0.94	0.21
3	Deisenhofen	0.51	0.64	1.90	1.07	0.21
4	Ettal	0.52	0.70	1.63	1.00	0.34
5	Garmisch-Partenkirchen	0.49	0.62	1.83	1.01	0.23
6	Gilching	0.48	0.76	1.88	0.96	0.25
7	Griesen	0.50	0.70	1.49	0.86	0.35
8	Halblech	0.53	0.71	1.96	1.03	0.21
9	Hindelang	0.45	0.57	2.09	1.16	0.18
10	Hohenpeissenberg	–	–	–	–	–
11	Kaufbeuren	0.55	0.74	1.28	0.84	0.35
12	Kochel	0.53	0.70	1.66	0.97	0.24
13	Kohlgrub, Bad	0.56	0.77	1.82	1.01	0.27
14	Kraftisried	0.52	0.68	1.97	1.07	018
15	Kreuth	0.42	0.65	2.07	1.18	0.17
16	Krün	0.49	0.70	1.96	1.10	0.34
17	Lenggries	0.47	0.66	1.85	1.06	0.29
18	Maisach	0.46	0.73	1.65	0.95	0.28
19	Marktoberdorf	0.59	0.73	1.93	1.14	0.23
20	München	0.51	0.61	1.62	0.99	0.19
21	Oberammergau	0.547	0.74	1.70	0.99	0.27
22	Oberschleissheim	0.48	0.72	2.00	1.05	0.19
23	Oy	0.48	0.60	1.72	1.08	0.25
24	Schwangau	0.56	0.65	1.57	0.91	0.30
25	Seeg	0.56	0.65	2.09	1.09	0.12
26	Schäftlarn	0.50	0.74	1.73	0.99	0.16
27	Steingaden	0.64	0.88	2.00	0.99	0.35
28	Schwaben	0.47	0.48	1.96	1.08	0.16
29	Schlehdorf	0.49	0.70	2.13	1.17	0.17
30	Vilgertshofen	0.52	0.68	1.93	1.02	0.22
31	Wielenbach	0.54	0.70	1.73	0.91	0.24

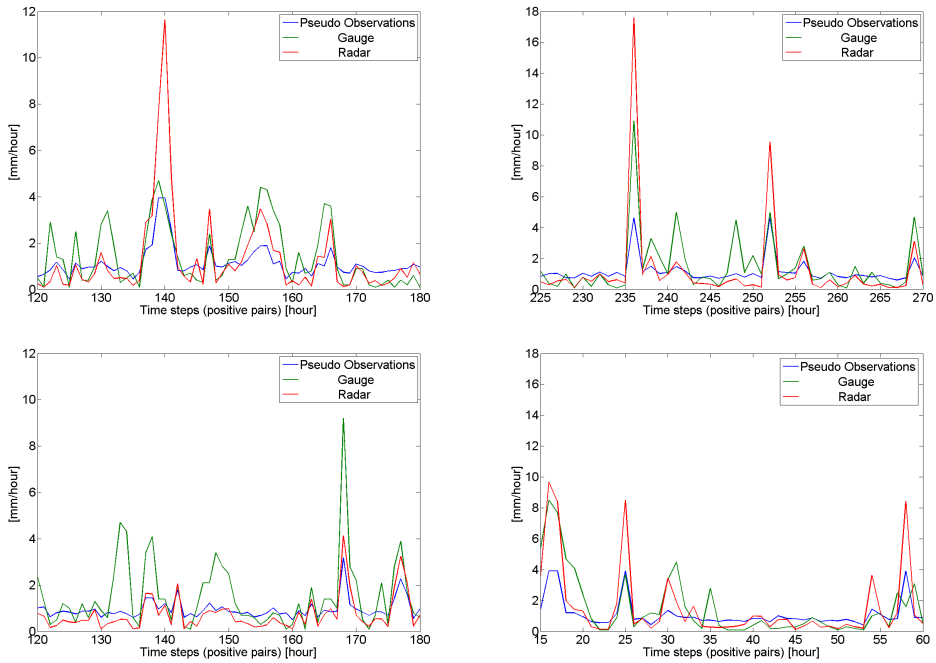

Figure 8.11: Time series of pseudo observations generated by using interpolated Copula parameter θ_G field based approach, the corresponding radar (Z/R-256/1.42) and gauge observations, for the station Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich (from left to right and top to bottom), summer, 2008, positive pairs only, w/ ARMA-GARCH transformation.

dependence structure can be defined for each grid cell in the study area. This interpolated Copula parameter θ_G field is good representation for the spatial distribution of gauge/radar dependence structure combining additional information from the range and altitude. The range is important for the radar measurements. With including altitude information, especially in the mountain area, the altitude dependent rainfall variations are also considered in this approach which is not included in the methods e.g. the Rodolan (DWD, 2001). For different seasons, the different interpolated Copula parameter θ_G field can be generated as shown in this chapter. In this way, the seasonal dependent circulation pattern and preferential rainfall types can be revealed. According to the GoF tes, the Gumbel Copula is used in this chapter. However, the other Copula family can also be easily integrated in this approach, as well as the different choices of rain gauge marginal distributions.

Then the pseudo observation fields can be simulated based on the interpolated

Copula parameter θ_G field. Comparing to the results from the *Multiple Theta* and *Maximum Theta* methods, the spatial pattern can also be captured in a reasonable way. Furthermore, the simulated pseudo observations from this new approach have a better merge with the real gauge observations, indicating the superior simulation results in both grid and spatial scale. The results w/o and w/ ARMA-GARCH transformation are also presented and compared in this chapter. The cross validation results also show that the ARMA-GARCH transformation leads to the worse simulations comparing to the real gauge observations.

To generate the interpolated Copula parameter θ_G field, the rain gauge located with high altitude (i.e. altitude larger than 1500 m) is still missing, which will directly influence the interpolation over altitude and range. On the other hand, this method is sensitive on the distribution of rain gauges, but only along the altitude and range to the radar station, not the absolute number and the homogeneous distribution for rain gauges in the whole area as the other methods such as *Multiple Theta* and *Maximum Theta* approaches.

Considering why to choose the bivariate polynomial in this study, the main reason is because of its simplicity and the convenient to be handle with. Of course, other theoretical function such as exponential function $f(Ra) = c \times exp(b \times Ra)$ can also be one option, which is used very often in the earth science to represent the geometry/spatial relationship. However, if the function $f(Ra)$ and $f(z)$ could be derivative in any order, according to the theory of Taylor series (this theory is a representation of a function as an infinite sum of terms that are calculated from the values of the function's derivatives at a single point), it is also practical to approximate a function by using a finite number of terms of its Taylor series. Taylor's theorem also gives quantitative estimates on the error in this approximation. Here, for example, the function $f(Ra) = c \times exp(b \times Ra)$ can be expanded to a Taylor polynomial like $exp(Ra) = 1 + Ra/1! + Ra^2/2! + Ra^3/3! + \cdots, here, b = 1$. As a result, if the truncation errors can be acceptable, the terms with high orders can be neglected so that the bivariate polynomial is also a reasonable way to represent the range and altitude related function $f(Ra)$ and $f(z)$.

Another problem that still remains is to find in the spatial distribution of marginal distributions for rain gauges. The same as the approaches in Chapter 7, for each grid, the fixed rain gauge marginal is selected from the existing 31 rain gauges. This problem will be further discussed in Chapter 9.

Nerveless, this newly proposed algorithm is a framework which can easily be applied to any test sites once the radar and gauge data are available. The computation cost is also very low and can be used as the real-time rainfall field estimation.

9

Combination of Various Spatial Distributions for Dependence Structure and Marginal Distribution

In Chapter 7 and 8, two different approaches to model the spatial distribution for radar/gauge dependence structure are developed, which can be referred to CPM and ICPF as listed below:

1. CPM ⇒ Copula Parameter Map based approach (*Maximum Theta*).
2. ICPF ⇒ Interpolated Copula Parameter Field based approach.

In addition to the radar/gauge dependence structure, the spatial distribution of gauge marginal distribution is the other key part of the Copula-based simulation technique. In previous chapters, only the fixed marginal distributions estimated from existing rain gauges are used to performed the back-transformation. This implies that the rain gauge marginal distribution is the same for each grid cells close to the existing rain gauge, no spatial variances included, which is not reasonable and realistic.

The structure of this chapter is as following. First, the newly developed approaches for rain gauge marginal distributions are introduced in section 9.1 and 9.2. After that, in section 9.3, simulation results, from the different combinations of spatial dependence structure and spatial gauge marginal distributions, are compared to the real gauge

observations in order to reveal the advantages and disadvantages of different methods. Then, the effect and impact from rainfall type classification and relative humidity are discussed in section 9.4 and 9.5, following the temperature dependent Copula in section 9.6.

9.1 Temperature and Altitude Driven Marginal Distribution

The estimation of gauge marginal distribution is one of the important steps in Copula-based analysis. In chapter 3, the univariate marginal have already been introduced and were also applied in Chapter 6 to estimate the marginal distribution both for radar and gauge positive observations. Afterwards, the impact from rain gauge marginal distribution on back-transformation to generate pseudo observations, has also been discussed in Chapter 6.

Note that in the previous chapters, the gauge marginal distributions, for those grids without rain gauges, are selected from the 31 existing gauges. This is achieved either by interpolating back-transformed values or just choosing the marginal distribution from nearest gauge, which is obviously not very reasonable and systematic. Therefore, this section focuses on developing systematic methods to model the spatial distribution of marginals for rain gauges.

The seasonal impact on the estimated parameters for gauge marginal distribution (Weibull distribution's scale and shape) has already been mentioned in chapter 6, as shown in Figure 6.3, implying the influences from temperature. In this section, the data set, from 20 DWD stations around the region of Bavaria with both precipitation and temperature observations for 15 years, are used to derive this temperature dependent marginal distribution for rain gauges.

9.1.1 Data Source

Data from 20 stations in Bavaria with both hourly rain gauge precipitation and temperature observations are used in this study during the period from 1995 to 2011. The geographic description of the 20 stations is listed in Table 9.1.

Available time periods are shown in Figure 9.1, for the 20 stations with both hourly precipitation and air temperature, Bavaria, Germany. Note that for station Grosser

Table 9.1: Description of the 20 stations with both hourly precipitation and air temperature observations, 1995 to 2011, Bavaria, Germany

ID	Station Name	Altitude [m.a.s.l]	Lat [°N]	Lon [°E]	∇ t [min]
1	Augsburg	461	48.43	10.93	60
2	Chieming	549	47.89	12.54	60
3	Fuerstenzell	476	48.55	13.36	60
4	Gebelsee	539	48.95	11.43	60
5	Grosser Arber	1446	49.11	13.14	60
6	Kempten	705	47.72	10.34	60
7	Muehldorf	405	48.29	12.51	60
8	Munich Airport	443	48.36	11.80	60
9	Munich City	535	48.14	11.55	60
10	Nuernberg	310.4	49.50	11.08	60
11	Obersdorf	806	47.40	10.28	60
12	Oehringen	276	49.21	9.52	60
13	Regensberg	366	49.04	12.10	60
14	Stoetten	733.8	48.67	9.87	60
15	Ulm	566.8	48.38	9.95	60
16	Weiden	438	49.67	12.19	60
17	Wuerzburg	268	49.77	9.96	60
18	Weissenburg	422	49.02	10.96	60
19	Garmisch-Partenkirchen	719	47.48	11.06	60
20	Hohenpeissenberg	977	47.80	11.01	60

Arber with the altitude at 1446 m, the data period is much shorter than others (only about 1/5 of the other). However, due to the high altitude (especially important for spatial extension), this station is still included here. Specific pre-processing procedures were performed, e.g. some unreasonable large values (as large as 100 mm/hour) are removed which just happened in 1995 due to the installation problem of these rain gauges.

9.1.2 Temperature Dependent Gauge Marginal Distribution

In order to explore the temperature dependent marginal distribution for rain gauges, the data from 20 stations listed in Table 9.1, both hourly precipitation and air temperature observations are used. As for one station, the first step is to separate gauge observed

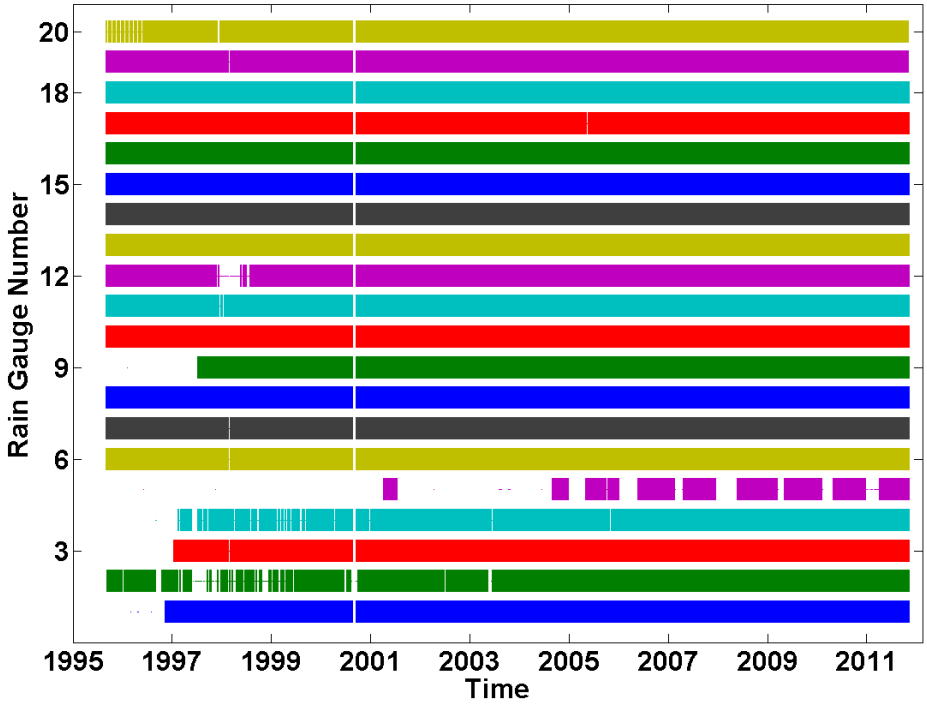

Figure 9.1: Available time period of 20 stations with both hourly precipitation and air temperature (arranged by the station ID number in Table 9.1), 1995 to 2011.

positive precipitation into different groups according to the corresponding temperature range (from 0 to 21 °C, 1 °C resolution). After that, the marginal distributions are fitted to each group.

To decide which univariate distribution fits best for the gauge observations in each group, both AIC/BIC GoF tests are used (see Chapter 3 and Chapter 6). Both the AIC/BIC results suggest that the Weibull distribution is the best fit one. However, the results for the Exponential distribution are very close as their difference is around 3 percent or even less (similar as the results listed in Table 6.1). Therefore, the one parameter Exponential distribution is used in this chapter for the simplification purpose, also with the advantage to be practical and easy to be handle with especially for the spatial extension. However, this is a disadvantage for this approach. Additionally, the problem for applying the Weibull distribution in this approach is presented and discussed in 9.1.4. In this chapter, the Exponential distribution is used to model the

rain gauge marginal distribution and the Weibull distribution is used for radar.

If the data number in one group is less than 50, the estimated parameter for marginal distribution in this group is set to NaN, because statistical analysis is meaningless for such a small data set.

Then, by assigning fitted parameters of the marginal distributions to the corresponding temperatures, the relationship between the estimated parameters of gauge marginal distribution and temperatures is extracted. This procedure is repeated for all the 20 stations. The basic steps to explore this temperature dependent gauge marginal distribution can be summarized as follows:

1. For one station with both hourly gauge and temperature observations, pre-processing is performed to exclude abnormal data.

2. The gauge data (only the positive observations) is classified into different groups according to the corresponding temperature.

3. For each group, the Exponential distribution is estimated to represent the marginal distribution of rain gauge observations.

4. The algorithm is repeated for all the 20 stations.

9.1.3 Point Scale Results

The results of the estimated parameters (Exponential distribution's inverse scale λ) for gauge marginal distributions, for different temperatures (0 to 21 °C, 1 °C resolution), are listed in Figure 9.2 at different stations, together with their corresponding average precipitation intensities in Figure 9.3.

It can be seen that, in general, for each gauge, the estimated parameters λ increase with the increasing temperatures T, which can be modelled by the linear interpolation $\lambda = aT + b$, with the parameter a and b.

However, as can also be seen in Figure 9.2, at the temperature above 17 °C (exact value dependent on different locations) - a turning point, the parameters λ tend to decrease for most cases. This may be due to the reasons listed as follows:

1. Limited number of observed rainfall events when the temperature $\geq$ 17 °C, with less than 1 percent of the total positive rainfall data numbers.

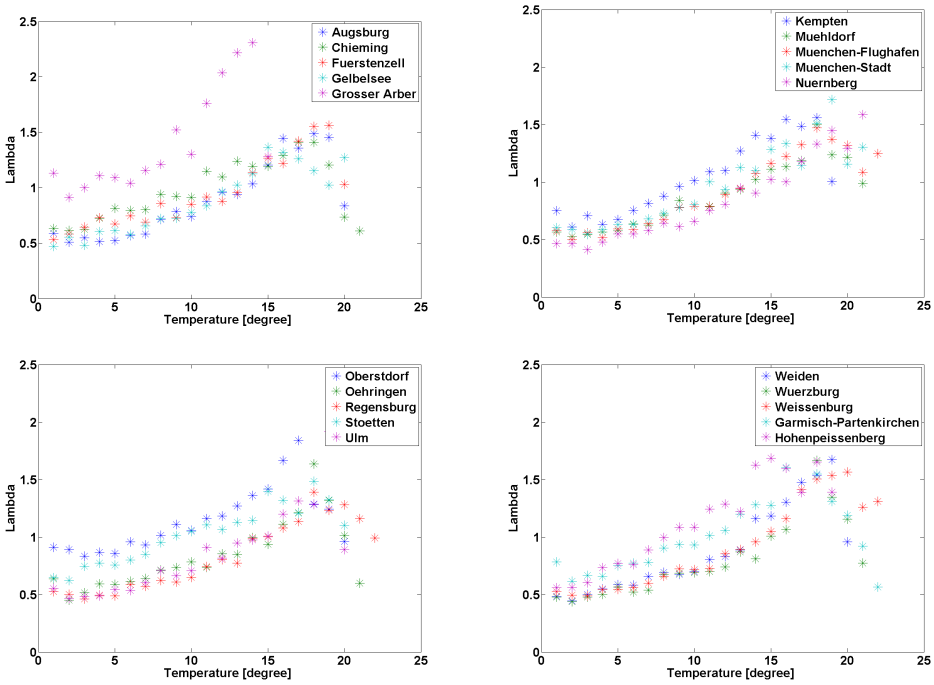

Figure 9.2: Parameter of the estimated Exponential distribution for different temperatures for all rain gauges, 1995 to 2011.

2. Higher probability for the occurrence of convective rainfall events when the temperature is larger than 17 °C.

Especially because of the first reason listed above, in this study, the focus lies mainly on temperatures below this turning point of 17 °C, which is also another limitation of this approach.

9.1.4 Spatial Extension

Considering the spatial extension of the temperature dependent gauge marginal distributions, the altitude information is included. The assumption is made that the parameters a and b (the linear interpolation $\lambda = aT + b$) are dependent on the corresponding altitude for each station. Once the temperature and altitude of one grid are known, the parameter λ for the corresponding gauge marginal distribution can be calculated. As a result, a virtual rain gauge can be assigned to each grid cell.

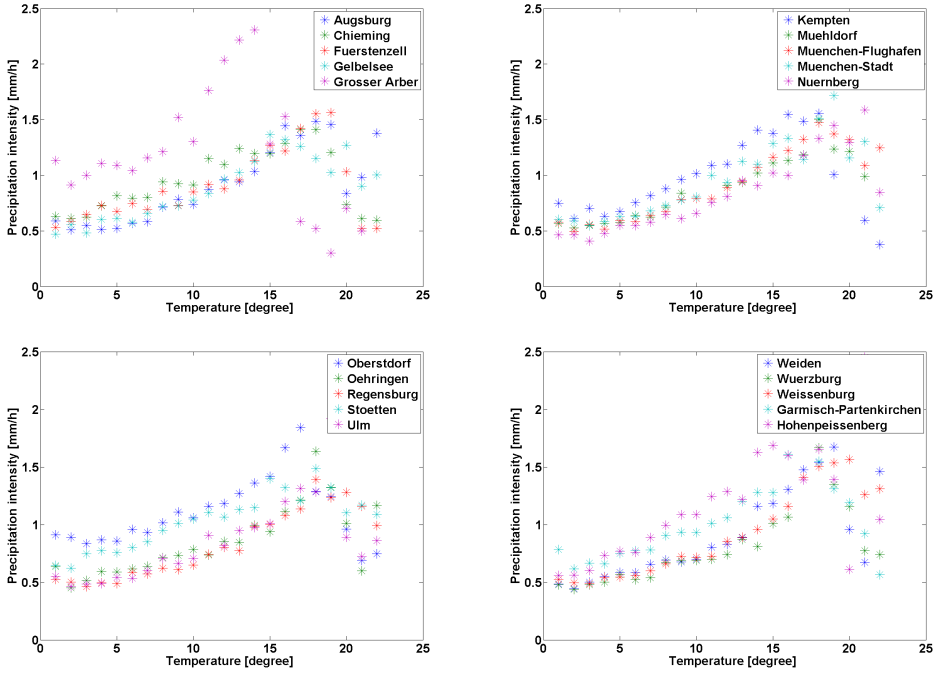

Figure 9.3: Mean precipitation intensity for different temperatures for all rain gauges, 1995 to 2011.

This altitude related assumption is also supported by the empirical results from those 20 stations as shown in Figure 9.4. Then, the following formulas, $f_a(z)$ and $f_b(z)$, are fitted to calculate the parameter a and b at different locations with the altitude z.

$$a = f_a(z) = 0.000027z + 0.026 \tag{9.1}$$

$$b = f_b(z) = 0.0004z + 0.26 \tag{9.2}$$

Eventually, the basic steps for this temperature and altitude driven rain gauge marginal distribution (Exponential distribution) can be summarized as follows:

1. For a specific radar grid cell, based on the altitude z, the parameters a and b (the linear interpolation $\lambda = aT + b$) are calculated by using $a = f_a(z)$ and $b = f_b(z)$.

2. Then, the parameter λ of virtual rain gauge marginal distribution (Exponential distribution) is calculated by using $\lambda = aT + b$, conditioned on the corresponding temperature T.

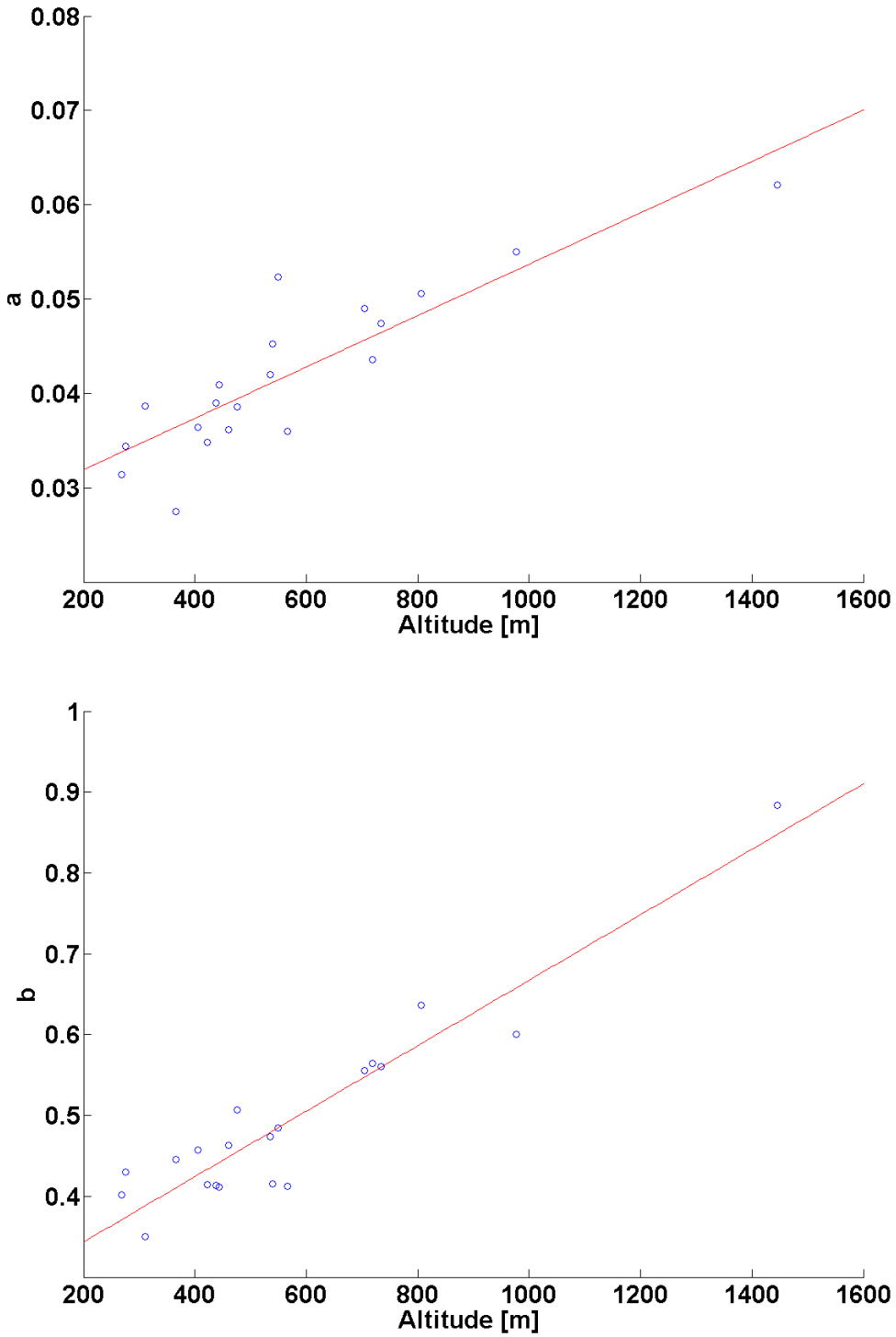

Figure 9.4: Linear interpolated coefficient a and b vs altitude (from top to bottom).

3. These steps are repeated for all grid cells to generate the parameter field of virtual rain gauge marginal distributions.

The example temperature and altitude driven rain gauge marginal distribution λ field for 13:00 14/07/2008 is shown in the top of Figure 9.5, together with the corresponding temperature field in the bottom. It can be seen that the temperature field is mainly dependent on the altitude, decreasing temperature with increasing altitude. The minimum temperature is achieved at Zugspitze with the highest altitude in Germany. Generally, the virtual rain gauge marginal λ shows a similar pattern as in the temperature field. However, at the regions with very high altitude, the impacts from temperature is stronger than altitude so that the small λ can be found in areas such as Zugspitze.

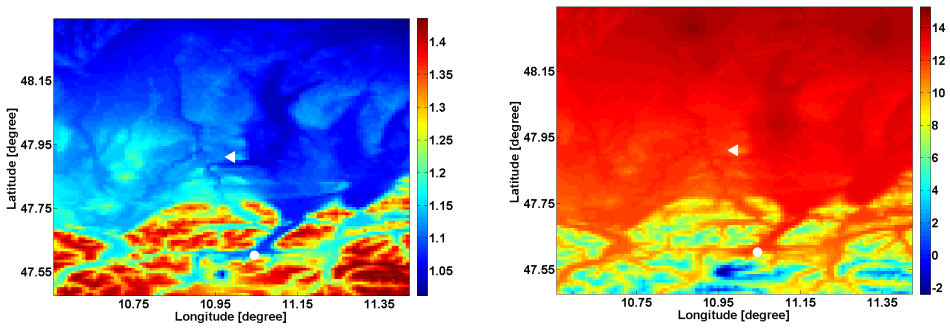

Figure 9.5: Simulated field of parameter λ derived from the temperature and altitude driven rain gauge marginal distribution (left) and the corresponding interpolated temperature field (right), 13:00, 14.07.2008.

9.1.5 Weibull Distribution Based Approach

In the previous sub-sections, although the Weibull distribution shows the best fit according to the GoF test, due to the purpose of simplification, the one parameter Exponential distribution is used to replace the Weibull distribution with two parameters. However, following the same method as introduced in 9.1.2, similar as for Exponential distribution, the temperature dependent pattern can still be found for the estimated scale and shape parameters of Weibull distribution for each station. The results at different locations are listed in Figure 9.6 and Figure 9.7, for scale and shape parameters separately.

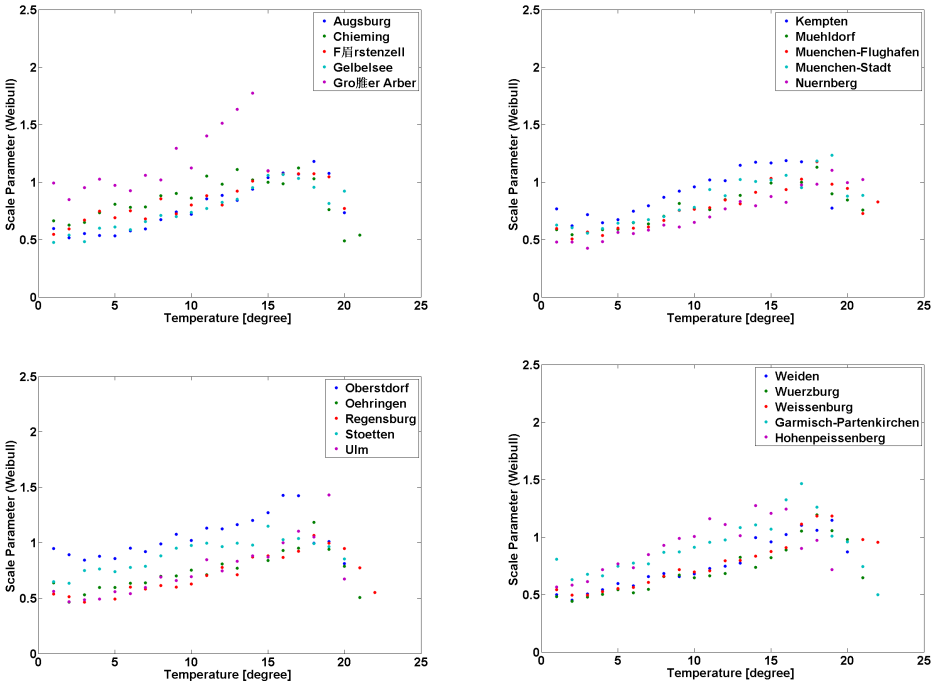

Figure 9.6: Scale parameter of the estimated Weibull distribution for different temperatures for all rain gauges, 1995 to 2011.

It is clear that, for all stations, the scale parameters increase with the increasing temperatures (before 17 °C) and the shape parameters decrease with the increasing temperatures (before 17 °C). As for the same reasons listed in 9.1.3, this study focus on temperatures below 17 °C.

The same as for Exponential distribution, for each station, the linear interpolation as listed below can also be applied to model the behaviour of the estimated scale P_{scale} and shape P_{shape} parameters of Weibull distributions for different temperatures.

$$P_{scale} = a_{scale}T + b_{scale} \tag{9.3}$$

$$P_{shape} = a_{shape}T + b_{shape} \tag{9.4}$$

Then, for the purpose of spatial extension, the estimated coefficients (a_{scale} and b_{scale}) for the scale parameters; a_{shape} and b_{shape} for the shape parameters) can be assigned to their altitudes at each grid cell where the gauge is located. The results are shown in Figure 9.8 and Figure 9.9. With the increasing altitudes z, the estimated

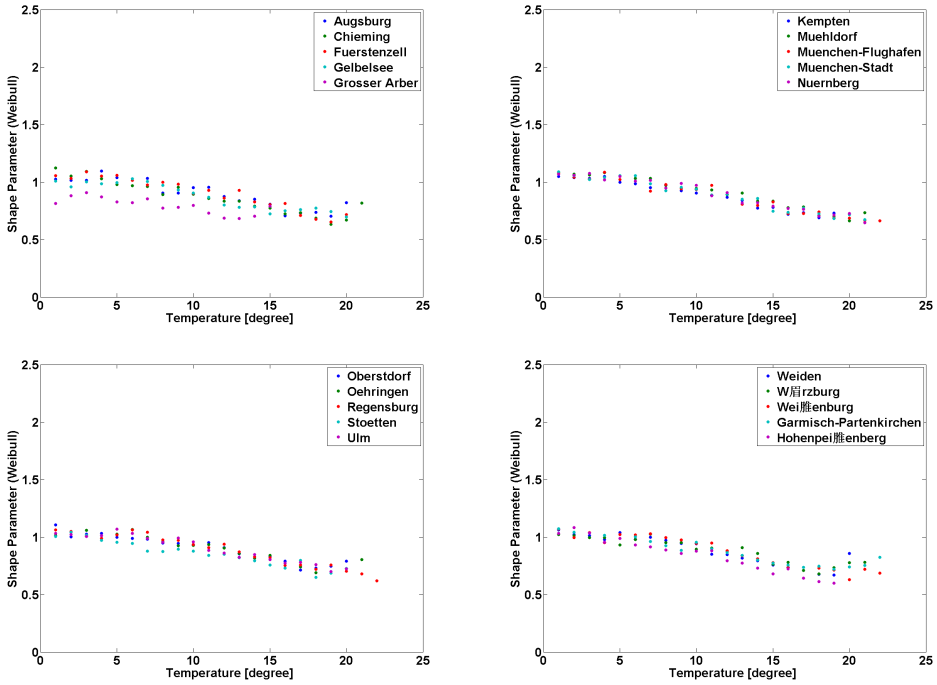

Figure 9.7: Shape parameter of the estimated Weibull distribution for different temperatures for all rain gauges, 1995 to 2011.

linear coefficients, a_{scale} and b_{scale} for scale parameters, all have a trend to increase so that a_{scale} and b_{scale} can also be calculated from the known altitude z by using the formulae listed below:

$$a_{scale} = c_{a1}z + c_{a2} \tag{9.5}$$

$$b_{scale} = c_{b1}z + c_{b2} \tag{9.6}$$

The estimated coefficients b_{shape} for shape parameters decrease with the increasing altitudes (at the bottom of Figure 9.9). However, the estimated coefficients a_{shape} for shape parameters are seemed to be randomly distributed along the different altitudes, mean value at around 0.02 and variation at around 0.005 (in the top of Figure 9.9). Then, the problem lies on to use certain theoretical function to model the relationship between a_{shape} and altitude z. If the small variation of the coefficients a_{shape} for shape parameters could be neglected, then it is also possible to perform the temperature and altitude driven marginal distribution for rain gauges by using the Weibull distribution with two parameters.

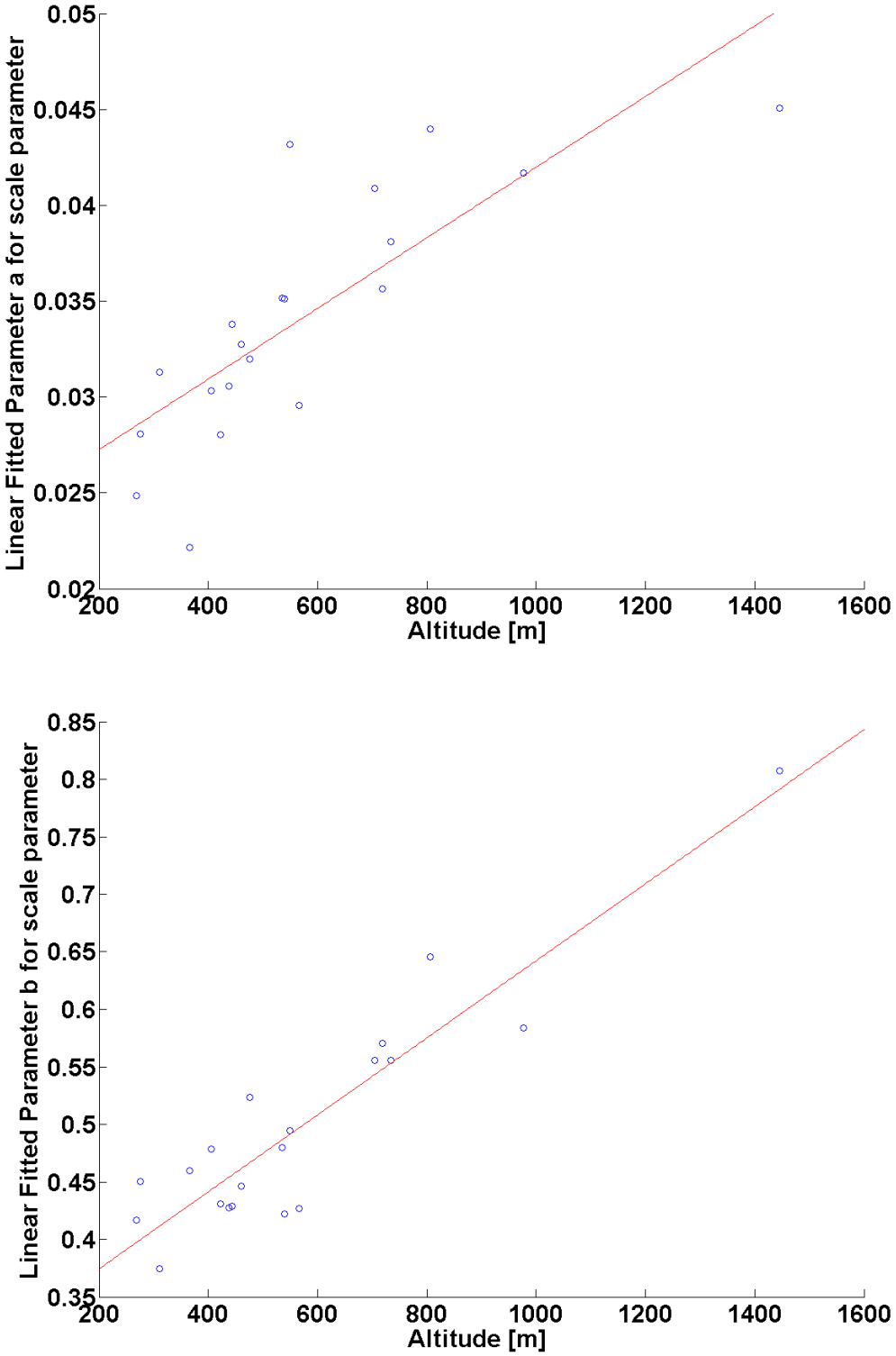

Figure 9.8: Linear interpolated coefficient a_{scale} and b_{scale} vs altitude (from top to bottom), Weibull distribution.

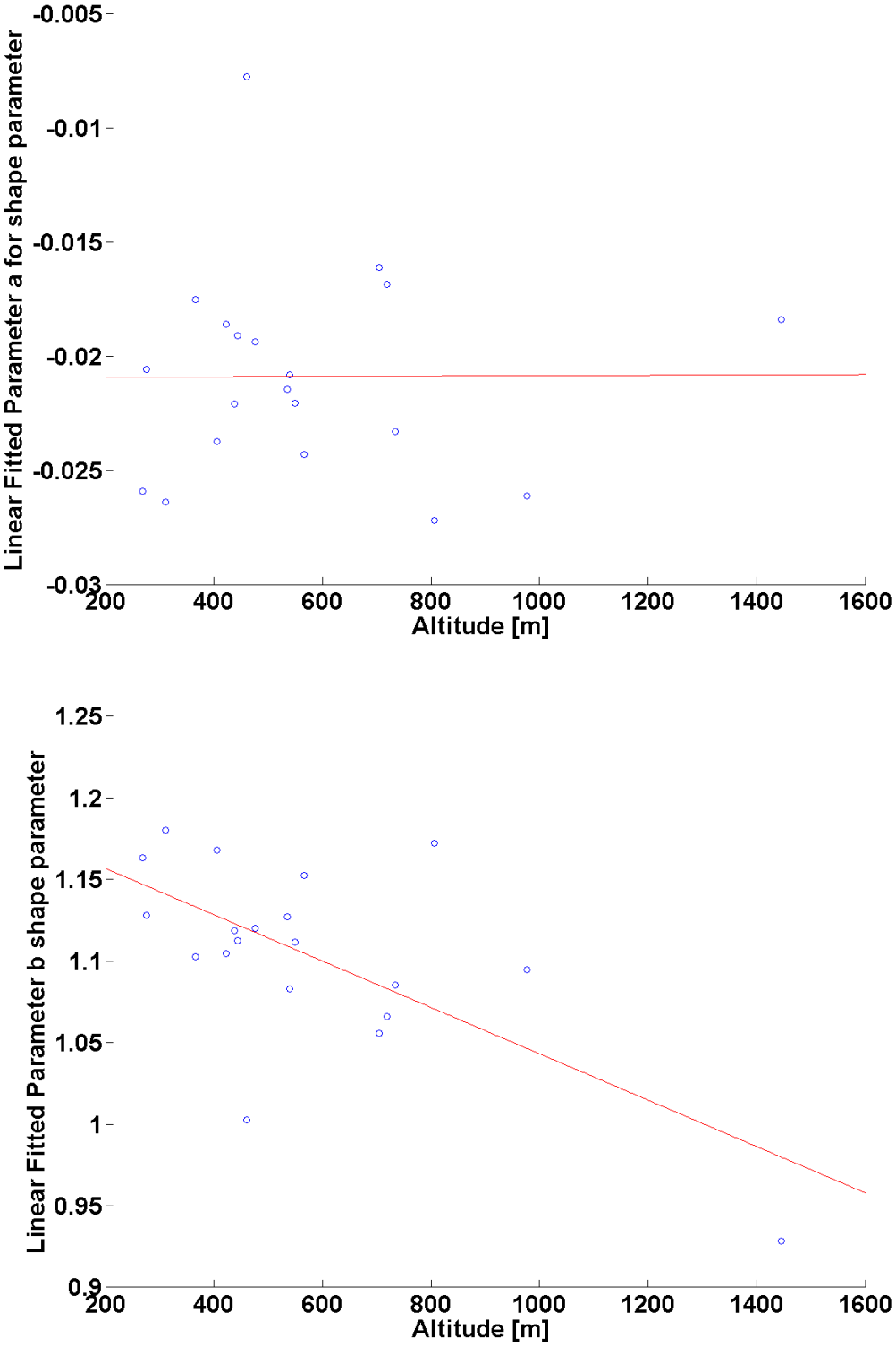

Figure 9.9: Linear interpolated coefficient a_{shape} and b_{shape} vs altitude (from top to bottom), Weibull distribution.

9.2 Rainfall Type Classification

It is known that the rainfall type classification is very important, due to the reason that different rainfall types indicating the physical principles behind. Since no radar vertical profiles are available for this study, the only chance to do the radar rainfall type classification is based on the horizontal structure of the radar measurement. Followed by the algorithm developed by Steiner et al. (1995), the basic steps to classify the convective and stratiform rainfall can be summarized as follows:

1. Intensity: any grid cell in the radar field with rainfall at least 13.4 mm/hour (40 dBZ) is automatically labelled as a convective center, since rain of this intensity should practically never be stratiform.

2. Peakedness: any grid cell not labelled as convective center after the firs step, but which exceeds the average intensity taken over the surrounding background by at least the intensity difference as shown in Steiner et al. (1995) is also identified as a convective center.

3. Surrounding area: For the grids cell identified as the convective center by one of the above two criteria, all surrounding grid cells within an intensity-dependent convective radius as shown in Steiner et al. (1995) around that the grid cell are also included as convective area.

Note that, in the work done by Steiner et al. (1995), the approach is performed on the radar reflectivities not on the rainfall intensities but there is a Z/R relationship between them. This is a limitation to apply this rainfall type classification in this study.

One example of the classified radar field is shown in Figure 9.10 also at 13:00, 04.08.2008. The convective regions are co-existing and surrounded by the stratiform rainfall cells. Because of the different micro-physical principles of the two rainfall types as introduced in Steiner et al. (1995), the statistical characteristics such as the marginal distributions and gauge/radar dependence structures are also different.

Performing the rainfall type classification methods for each time steps, during summer from 2006 to 2007, the estimated marginal distributions in different rainfall types are shown in Table 9.2, as well as the corresponding Copula parameter. Generally, there are huge differences between the estimated marginal parameters in the two rainfall types both for gauge and radar, as high as 200 percent or even more. In the sense

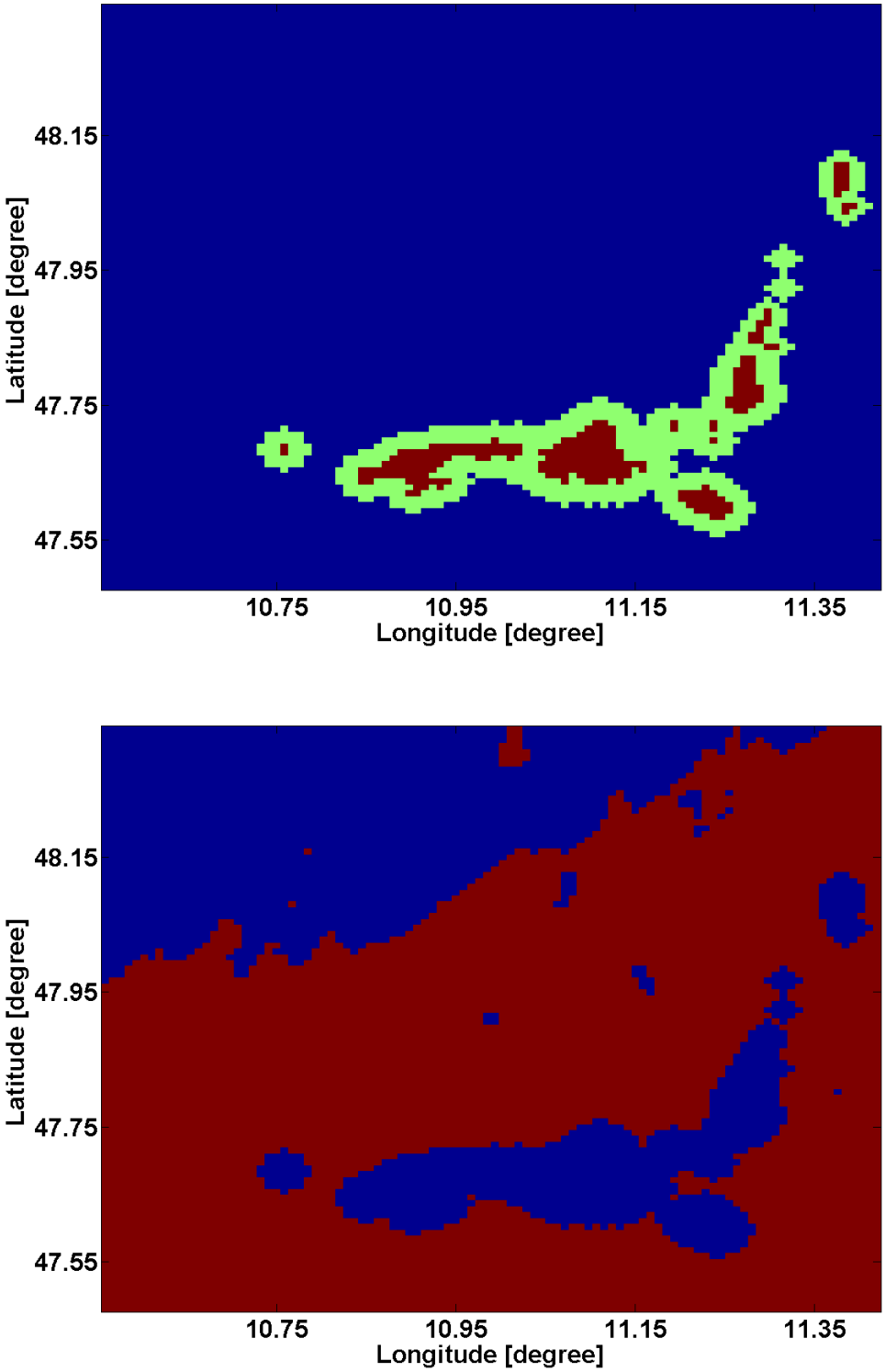

Figure 9.10: The radar field with convective rainfall (top, highlighted as brown) and stratiform rainfall (bottom, highlighted as brown in the first step and light green in the second step) at 13:00, 04.08.2008

of Copula parameters, the difference is not as large as for the marginal distributions, but still around 30 percent.

Table 9.2: Comparison of marginal distributions for gauge/radar (parameter λ of Exponential distribution), and Copula parameter (θ_G) estimated in convective and stratiform rainfall types for Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich, positive pairs, summer, 2006 to 2007, w/o ARMA-GARCH transformation.

	Garmisch-P.		Oberammergau	
	Convective	Stratiform	Convective	Stratiform
Gauge Marginal λ	3.43	1.24	3.16	1.35
Radar Marginal λ	4.86	1.21	5.20	1.50
Copula Parameter θ_G	1.86	2.01	1.76	2.41
	Wielenbach		Munich	
	Convective	Stratiform	Convective	Stratiform
Gauge Marginal λ	2.87	1.28	3.31	1.41
Radar Marginal λ	4.57	1.32	5.31	1.61
Copula Parameter θ_G	1.63	2.28	1.45	1.94

These classified marginal distributions are supposed to overcome the problem mentioned in Chapter 6 leading to large rank value also with large value in data space, as the gauge marginal distributions for stratiform rainfall is similar as that without classification and in convective case is much larger. However, during summer 2006 to 2007, the number of convective rainfall events is very limited (only less than 5 or 6 percent in the total rainfall events). This is another disadvantage for the application of this rainfall classification approach in this study, as can also been seen for the simulation results listed in the next section.

9.3 Combinations of Different Spatial Dependence Structure and Marginal Distributions

Until now, three different approaches to model spatial marginal distributions for rain gauges are already developed:

1. FMD ⇒ Fixed Marginal Distribution estimated from existing rain gauges.
2. TAMD ⇒ Temperature and Altitude driven Marginal Distributions for rain gauges.
3. CSMD ⇒ Convective and Stratiform dependent Marginal Distributions for rain gauges.

So, in this section, the combinations of different spatial dependence structures and gauge marginal distributions are tested and compared to the corresponding gauge observed positive series at summer 2008 by using the point wise cross-validations. The purpose of this part is to study the performances of different combinations of spatial approaches for dependence structures and gauge marginal distributions. Note that, in this chapter, the Gumbel Copula is used for all dependence structure, Exponential distribution for all rain gauge marginal distributions and Weibull distribution for all radar marginal distributions. Furthermore, the ARMA-GARCH transformation is not considered in this chapter as for the reason of simplification. All the 6 possible combinations of the 2 approaches for spatial dependence structures and 3 approaches for rain gauge marginal distributions are listed as follows:

1. CPM + FMD
2. CPM + TAMD
3. CPM + CSMD
4. ICPF + FMD
5. ICPF + TAMD
6. ICPF + CSMD

The point wise cross-validation results, obtained from these 6 different combinations, are listed in the Table 9.3, Table 9.4 and Table 9.5, separately for the different indexes of Pearson's correlation coefficient, RMSE and NSE, at Garmisch-Partenkirchen, Oberammergau, Wielenbach and Munich, summer, 2008.

For the 4 different locations, the mean values of Pearson's correlation coefficients are around 0.66, 0.76, 0.72 and 0.66, all with small deviations and therefore showing equally good correlation structures between all the Copula-based simulations and gauge

Table 9.3: Pearson's correlation coefficient calculated between rain gauge observations and generated pseudo observations from different combination of spatial distribution for the dependence structure and gauge marginal distributions, positive pairs only, summer, 2008, w/o ARMA-GARCH transformation.

Combination ID	Garmisch-P.	Oberammergau	Wielenbach	Munich
CPM+FMD	0.65	**0.78**	0.72	0.65
CPM+TAMD	**0.67**	0.76	0.72	0.66
CPM+CSMD	0.65	0.77	0.72	0.65
ICPF+FMD	0.65	**0.78**	**0.73**	**0.68**
ICPF+TAMD	0.66	0.73	0.69	0.62
ICPF+CSMD	0.65	0.77	0.72	0.67

observed series (positive pairs, excluded zero precipitation). The reason is that the Pearson's correlation coefficient in the cross-validation is mainly dependent on the given radar measurements. For different locations, the best simulation (highlighted in bold for each location) is achieved by using the different combinations, showing the spatial variations. At Garmisch-Partenkirchen and Oberammergau, close to or within the mountains, the CPM performs a little bit better or equally good as the ICPF. At Wielenbach and Munich, the flat and city area, the ICPF+FMD shows the best results.

Table 9.4: RMSE [mm/hour] calculated between rain gauge observations and generated pseudo observations from different combination of spatial distribution for the dependence structure and gauge marginal distributions, positive pairs only, summer, 2008, w/o ARMA-GARCH transformation.

Combination ID	Garmisch-P.	Oberammergau	Wielenbach	Munich
CPM+FMD	1.81	1.63	1.69	1.53
CPM+TAMD	**1.63**	**1.48**	**1.65**	**1.52**
CPM+CSMD	1.76	1.66	1.69	**1.52**
ICPF+FMD	1.81	1.66	**1.65**	**1.52**
ICPF+TAMD	1.73	1.59	1.71	1.60
ICPF+CSMD	1.76	1.69	1.66	1.53

Considering about RMSE and NSE, the combination of CPM+TAME shows the best results at Garmisch-Partenkirchen and Oberammergau, the mountain area, 12 percent better than the other combinations of approaches. At Wielenbach, the flat area, the ICPF+FMD and CPM+TAME show the same best results (RMSE) or very close results (NSE). At Munich within the city area, the situation becomes complicated, the best results are achieved by CPM+TAMD, CPM+CSMD and ICPF+FMD for RMSE; by CPM+TAMD, ICPF+FMD and ICPF+CSMD for NSE;

It is worth to note that, the results from CPM+TAMD (Exponential) are also better than the results by using CPM+FMD (Weibull distribution) as listed in Table 7.2. That may be due to the reason that, for the purely statistical Copula parameter map based approach, the temperature and altitude driven rain gauge marginal distributions can be a good complement because the physical background can be added so that the final results are improved.

In contrast, the interpolated Copula parameter field based approach, the altitude information is already included in the spatial distribution for the dependence structure. Therefore, in this case, the temperature and altitude dependent rain gauge marginal distributions may have some unexpected problems because the altitude information is used for two times, both for Copula and marginal distribution. However, there are other problems for temperature and altitude dependent rain gauge marginal distributions and a more detail discussion can be found in the next section.

Table 9.5: NSE calculated between rain gauge observations and generated pseudo observations from different combination of spatial distribution for the dependence structure and gauge marginal distributions, positive pairs only, summer, 2008, w/o ARMA-GARCH transformation.

Combination ID	Garmisch-P.	Oberammergau	Wielenbach	Munich
CPM+FMD	0.25	0.32	0.27	0.28
CPM+TAMD	**0.39**	**0.44**	0.31	**0.29**
CPM+CSMD	0.29	0.30	0.28	0.28
ICPF+FMD	0.25	0.30	**0.32**	**0.29**
ICPF+TAMD	0.31	0.35	0.26	0.21
ICPF+CSMD	0.29	0.27	0.31	**0.29**

As listed in the Table 9.3, Table 9.4 and Table 9.5, the simulation results, by performing the rainfall type classification, are not improved but become even a litter bit worse. Some possible reasons are already given in 9.2 as the limitations for this rainfall type classification to be applied in this study. Another reason is that this rainfall type classification is only based on the horizontal radar field and no vertical profile available in this study. This will also have unexpected influences on the final results.

9.4 Impacts from Relative Humidity Effect for Marginal Distributions

It has been proven that the gauge marginal distribution is dependent on the temperature in section 9.2 and the temperature information was also included to generate the simulation fields of pseudo observations. However, the simulation results are not very promising as listed in the last section. Further analysis reveals that another variable - relative humidity - can also have influences on the gauge marginal distributions. By applying the same method as introduced in section 9.2, the gauge precipitation can also be separated into different groups according to the observations of relative humidity (from 70 to 100 percent, 1 percent resolution). The fitted gauge marginal distributions (parameter λ of Exponential distribution) are also assigned to the corresponding relative humidity ranges as shown in the top of Figure 9.11 at Garmisch-Partenkirchen from 1995 to 2011. This parameter λ also increases with the increasing relative humidity. When the relative humidity is above 97 percent, the estimated parameters become to be unclear. This is also mainly due to the reason that the rainfall data number is limited for the case with the relative humidity larger than 97 percent.

Furthermore, the behaviour of rain gauge marginal distributions for different temperatures and relative humidities is presented in the bottom of Figure 9.11 also at Garmisch-Partenkirchen from 1995 to 2011, illustrating a much more sophisticated structure of the rain gauge marginal distribution. There is a data availability problem as only 7 stations with relative humidity observation can be found in this study area. Therefore, it is difficulty to calculate the relative humidity field, also due to its physical background. This also implies that the spatial extension of temperature and relative humidity dependent gauge marginal distributions is still a challenge.

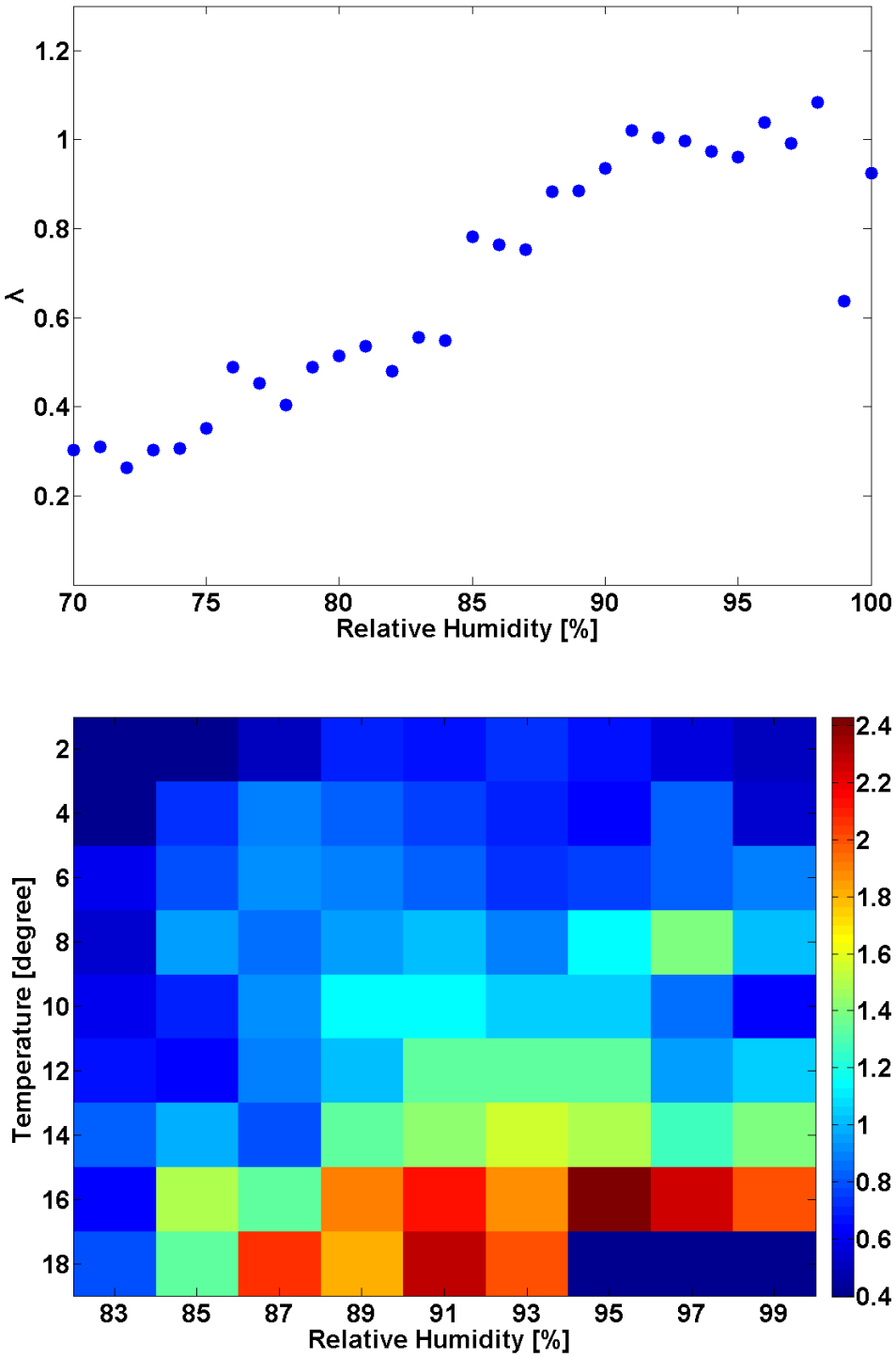

Figure 9.11: Parameter of the estimated Exponential distribution (rain gauge) for different different values of relative humidity (top) and both temperature/relative humidity (bottom), Garmisch-Partenkirchen, positive observations only, 1995 to 2011.

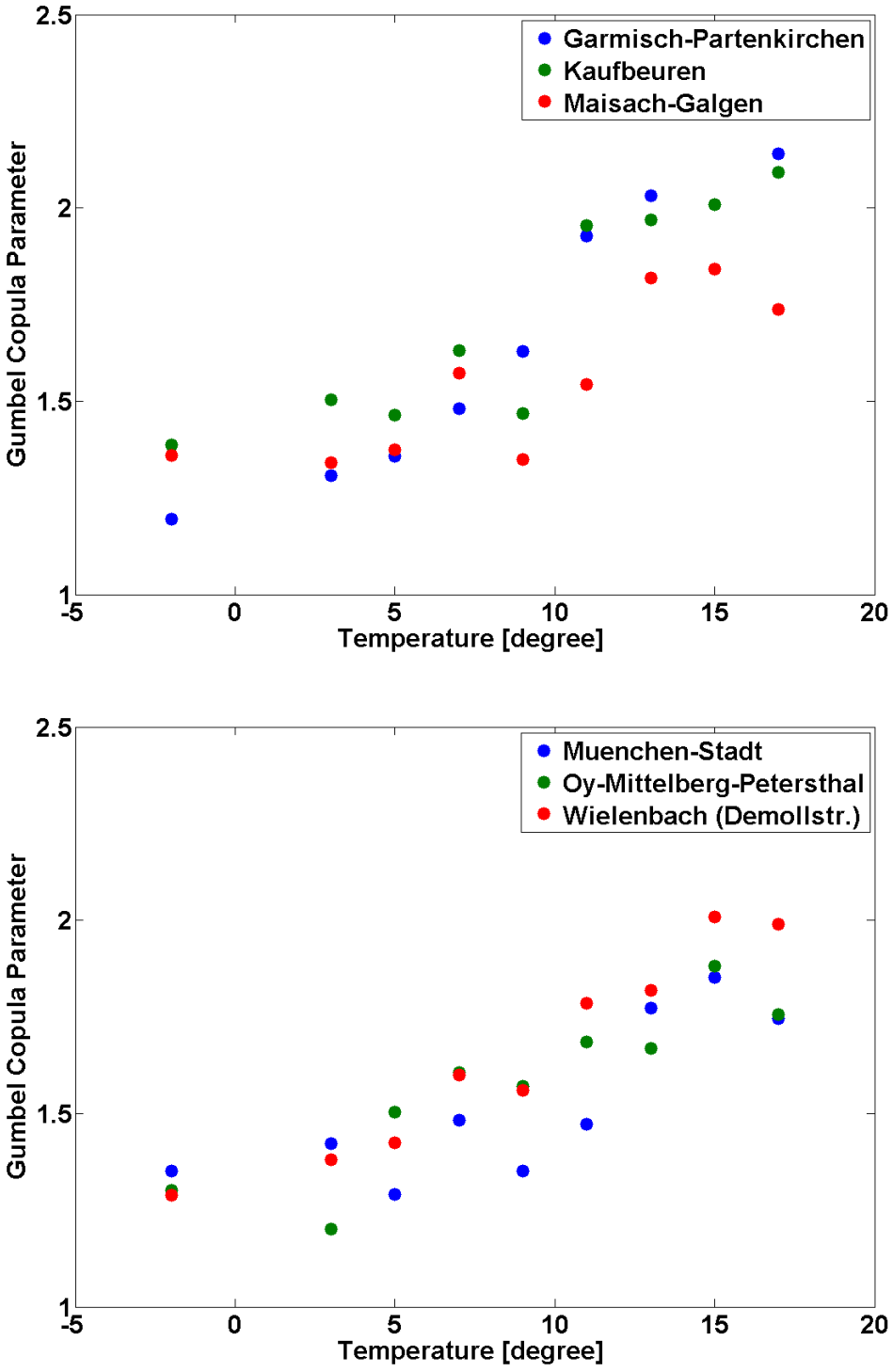

Figure 9.12: Estimated Copula parameter θ_G for different temperatures at 6 locations with all the radar, gauge and temperature observations in this study area, positive pairs only, 2005 to 2008, w/o ARMA-GARCH transformation.

9.5 Temperature Driven Dependence Structure

The temperature has not only an impact on gauge marginal distributions but also on the gauge/radar dependence structure - i.e. the Copula parameter θ_{G}. The same classification, as described in the section 9.2, is also applied to the 6 grid cells with all the radar, gauge and temperature observations during the time period from 2005 to 2008. Then empirical relationship, between the estimated Copula θ_{G} and the corresponding temperature range (-5 to 20 °C, 2 degree resolution), is shown in Figure 9.12 for all the 6 locations.

It can be seen that the Copula parameter θ_{G} increases with increasing temperature. However, when the temperature is above 17 or 18 °C, the trend of decreasing (not for the gauges located at Garmisch-Partenkirchen and Kaufbeuren) can also be found as for the temperature dependent gauge marginal distributions. Although it is not clear that this is due to the limited data number available for high temperature values or not, the explanation for this temperature dependent Copula θ_{G} may be the same as listed in 9.2 for the temperature dependent gauge marginal distributions.

In this study area, only for 6 stations, all the measurements from radar, gauge and temperature are available. Consequently, it is not easy to do the spatially extension similar as for temperature and altitude dependent gauge marginal distributions. However, another approach can be considered, as can be seen from the results listed in Figure 9.12. The exact parameters of linear fitting can be different for different locations. However, the ΔT and $\Delta\theta_G$ show only very small variations or nearly the same. However, this still needs more efforts. More data is required to complete the spatial extension for temperature driven gauge/radar dependence structure.

9.6 Summary and Discussion

In this chapter, at first, the two new developed spatial distributions for rain gauge marginal were introduced. Then the simulation results, from different combinations of the dependence structures and spatial rain gauge marginal distributions, were also tested by using the point wise cross-validation. At Wielenbach and Munichn, the flat or the city area, the comparison results show almost equally good performances for all the combinations. It is found that, the Copula parameter map based approach can be clearly improved by using the temperature and altitude driven rain gauge marginal

distributions and shows the best results, especially in the locations within or close to the mountains such as Garmisch-Partenkirchen and Oberammergau.

According to the GoF test, the Weibull distribution is the best fit one for the rain gauge marginal distribution, which is main the disadvantage to use Exponential distribution in the temperature and altitude driven marginal distributions. However, by using the temperature and altitude driven marginal distributions, the *Maximum Theta* approach can perform even better than using the fixed Weibull distribution for rain gauge marginal distributions. For each station, the temperature dependent pattern can also be found for both the scale and shape parameters of Weibull distribution, similar as the Exponential distribution. However, the problem still remains for the spatial extension for the Weibull distribution.

However, there are some problems existing for the rainfall type classification approach and temperature/altitude dependent rain gauge marginal which are further detailed discussed in section 9.4 and 9.5. The impact for gauge marginal distribution from relative humidity is studied. Then, in this way, the gauge marginal distributions are directly related to both the temperature and relative humidity. Furthermore, if the air temperature and relative humidity data with small resolution (i.e. 5 min), it is also possible to determine the occurrence of precipitation (wet/dry identification) by statistical analysis results from those data sets only. However, due to the difficulty to simulate the field for relative humidity, the spatial extension, for the temperature and relative humidity driven gauge marginal distribution field, still needs larger observations and more efforts.

The influence of temperature on the gauge/radar dependence structure is also introduced in this chapter. In this way, for different temperatures, different Copula parameters can be generated, not like the fixed one used in this study. This may be one of the potential ways to realise an alternative version of time varying or dynamic Copula approach. However, due to the limited number of grid cells with all radar, gauge and temperature observations, the spatial extension for this relationship is still a challenge and waiting to be solved.

10

Simulation Including MW-Links

10.1 Introduction

MW-link attenuation is an independent and complementary device comparing to gauge and radar, as already introduced in Chapter 5. It has been proved that the rainfall derived from those wireless communication networks is qualified at grid scale. In order to estimate the spatial precipitation field, it is straightforward to separate the line integrated rainfall from MW-link attenuation and many efforts have been done in the past. The tomographic reconstruction (e.g. Giuli et al., 1991, 1999 and 2007) was introduced to estimate the rainfall fields by using MW-link attenuation. However, this approach required a specially designed hypothesized system of MW-links with a predefined geometry, operating at specially selected frequencies where the A/R relationship is linear, combined with point rain gauge observations. Another tomography based interpolation method (Zinevich et al., 2008) was developed to separate the MW-link observations into rainfall intensities for grid cells beneath this MW-link, and then to estimate the rainfall field by using interpolation approach such as Kriging. However, the application for this method is also dependent the density of the MW-links and therefore has almost no chance to be used in regions with poor density of the MW-link distribution. Besides from tomography, another method by using MW-links to adjust the radar rainfall field was given e.g. in Cummings and Upton (1996). In this method, a mean field bias adjustment is presented that uses the path-integrated rainfall estimations provided by MW-links together with information from gauges, similar as traditional approaches to adjust radar fields by using the gauges.

Due to the length and altitude gradients of those commercial MW-links, not only precipitation informations but also the impacts from other atmosphere variables such as temperature and relative humidity are included in the attenuation from the MW-links. As a result, the challenge still lies on how to interpret and make use of different information integrated in the measurements of MW-links. In this chapter, the Copula-based data assimilation approaches are further applied to integrate precipitation information from MW-link attenuation.

The structure of this chapter is as following. First in section 10.2, the Copula parameter maps derived from MW-link attenuation are described and then applied to produce the precipitation fields of pseudo observations, together with the validation procedures. In section 10.3, an approach to adjust radar observed rainfall field is introduced to use the MW-links as conditional and independent constraints. Finally, this chapter ends with the summary and discussion in section 10.4.

10.2 Copula Parameter Map Based Approach

The Copula parameter map, generated between MW-link and radar field, is also the basis to include precipitation information from MW-link. Then the simulated field of pseudo observation is presented and compared to results by using gauge/radar only. Afterwards, the validation is performed to test the quality of the simulations based on all the three data sources. Finally, the conclusion is given. Note that, due to the limited observation time period, the ARMA-GARCH transformation is not applied in this chapter.

10.2.1 Copula Parameter Map Derived from MW-Links

Similar as the approach introduced in Chapter 7, through treating the rainfall observations from the MW-link as point equivalent measurements (for instance to assume the MW-link is centred at its middle), the Copula parameter maps (Gumbel Copula, according to the GoF test resutls similar as listed in Chapter 6 for radar and gauge) can also be calculated between this MW-link and the whole radar field. The basic steps to calculate the MW-link derived Copula parameter map can be summarized as follows:

1. For all MW-links $L_i, i = 1, ..., n_L$ in the study area and all the radar pixels $R_j, j = 1, ..., m$, during certain training period, the corresponding marginal distributions are estimated.

2. For one specific MW-link, as well as all the m radar pixels, the corresponding Copula parameters θ_G for these m radar/MW-link positive pairs are calculated.

3. This procedure is repeated for each existing MW-link in the whole domain.

4. Finally, n_L Copula parameter maps can be derived between n_L MW-links and radar fields

As shown in Figure 10.1, the two Copula parameter maps are calculated from two MW-links, Hohenpeissenberg to Weilheim at the top and Hohenpeissenberg to Murnau at the bottom. Compared to the Copula parameter maps derived from gauges as shown in Figures 7.2, these MW-link derived Copula parameter maps cover larger areas with higher Copula parameters, implying a stronger dependence structure between MW-link and radar. Of course, with the increasing distance from the MW-link, the dependence structure also decreases. The asymmetric structure can also be found in the MW-link derived Copula parameter maps, but not very strong due to the reason that the altitude gradients are not very high in this region. Nerveless, the anisotropic nature of the dependence structure has also be considered and modelled in the Copula parameter maps derived from MW-links. The interesting point is that the strongest dependence structures in those two MW-link derived Copula parameter maps are achieved at the pixels a little bit distant away from the MW-links, not at the grids where beneath the MW-links.

The MW-link derived Copula parameter map can be considered as the combinations/integration of several Copula parameter maps calculated from gauges located beneath this MW-link. In fact, one disadvantage for this approach is to treat the MW-link as an equivalent point scale measurement. However, through the Copula parameter map, the integrated information contained in the MW-links can still be extracted by using the spatial distributed radar field.

Note that Figure 10.1 only shows the areas close to the radar station in Hohenpeissenberg, that is because of the radar data availability problem from June to October, 2010. Due to the lower elevation angles setted for this C-band experimental weather

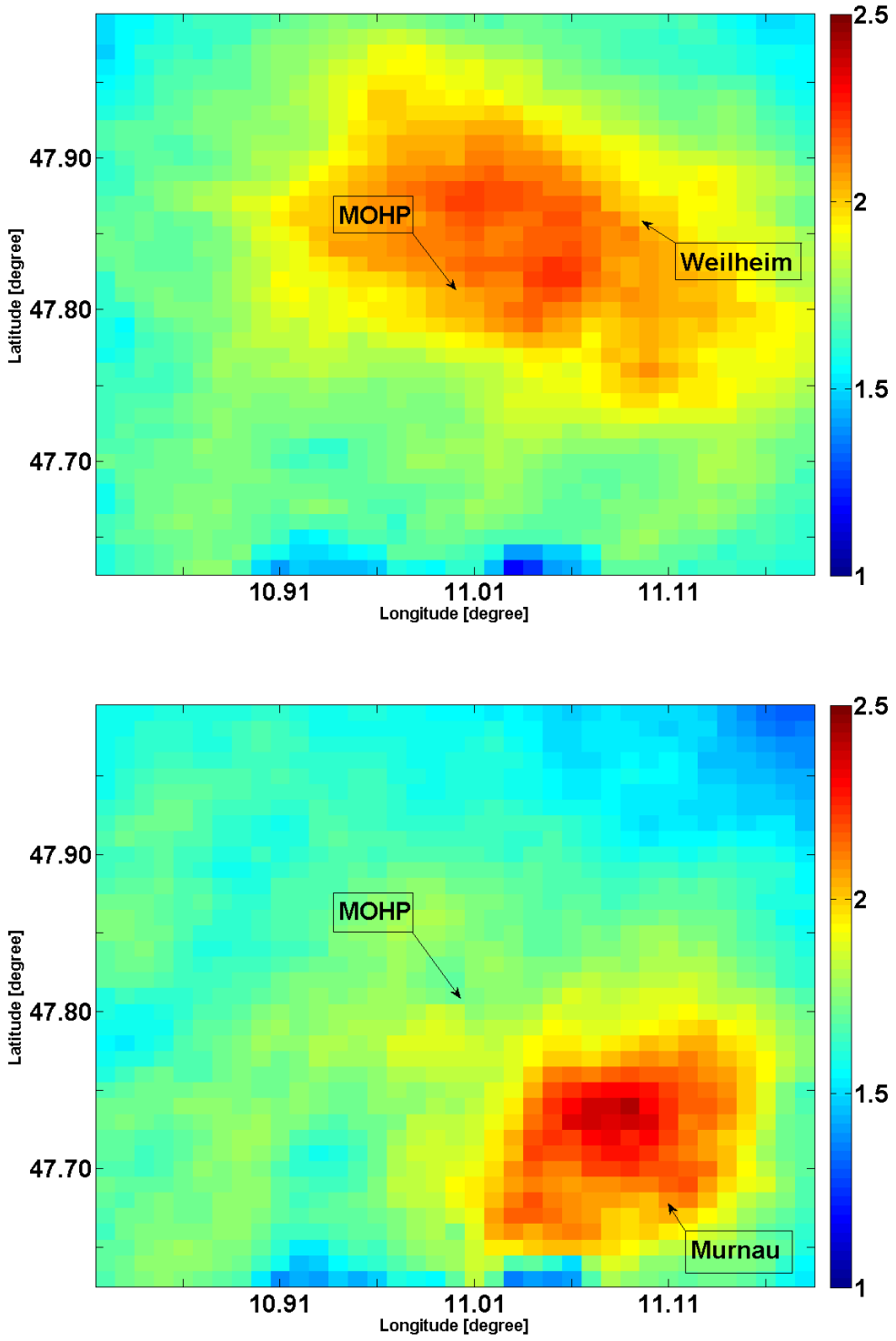

Figure 10.1: Copula parameter maps (Gumbel) derived from MW-link 1 (Hohenpeissenberg to Weilheim, top) and MW-link 2 (from Hohenpeissenberg to Murnau, bottom), positive pairs, June to October, 2010. The color bar is the value of the Copula parameter θ_G.

radar at Hohenpeissenberg, the radar rays are totally blocked by the mountains around Oberammergau so that no radar measurements for areas with the latitude below around 47.60 °N. Therefore, the Copula parameter maps can not be generated for the 3 MW-links around Garmisch-Partenkirchen. Nerveless, in the following sub-sections, these two MW-link derived Copula parameter maps are integrated in the *Maximum Theta* approach to produce the precipitation fields of pseudo observations combining rainfall observations from radar, gauge and MW-link.

10.2.2 Simulated Field of Pseudo Observations

Since the MW-link is treated as an equivalent point scale device, the assumption can be made that those MW-link derived Copula parameter maps are equivalent to their counterparts generated from the rain gauges. Therefore, together with all the available Copula parameter maps from gauges, the MW-link derived Copula parameter maps are integrated in the *Maximum Theta* approach to generate the simulated field of pseudo observations conditioned on the given radar observed rainfall field for each time step. Note that, for the back transformation to the data space, the marginal distributions are selected from the gauges with maximum value among all the gauge derived Copula parameter maps for each grid cell. Similar as already described in Chapter 7, the basic steps can be summarized as follows:

1. Based on the available n_G gauge stations and n_L MW-links, $n_G + n_L$ Copula parameter maps are prepared, resulting in a set of $n_G + n_L$ Copula parameters for a specific radar grid cell.

2. The set showing the maximum Copula parameter is assigned for a specific grid cell, retaining the information of the Copula parameter and the corresponding marginal distribution (rain gauge).

3. One sample of 100 members is simulated in the rank space conditioned on the given radar measurement by using the maximum Copula parameter and the mean value is calculated from the random samples.

4. The integral transformation is applied to the mean value to transform back to the data space by using the selected rain gauge marginal distribution.

5. These steps are repeated for all radar grid cells to generated the precipitation field of pseudo observations.

At the test time step - 06:00, 04.08.2007, by applying this *Maximum Theta* approach, the simulated field of pseudo observation is shown in the middle of Figure 10.2 by using gauge/radar/MW-link, as well as the simulation only based on gauge/radar and original radar field at the top and bottom. Considering the whole domain, the basic precipitation spatial pattern from radar map is kept in a reasonable way in the simulated field of pseudo observations and the bias of radar measurements is corrected. Note that, also because of the data availability problem (only around 3 months in 2010 for MW-links and limited spatial coverage of radar data), the test period is chosen at summer, 2008(2007), which is the limitation to apply Copula-based approaches.

It can be also seen that, from Figure 10.2, in that areas highlighted with white ellipses, some more detailed patterns can be extracted by integrating precipitation information from MW-links compared to the result only based on gauge/radar. As shown in Figure 10.3, the area covered with the Copula parameters from the 2 MW-link derived Copula parameter maps is highlighted with red, which is nearly the same area within the white ellipses in Figure 10.2. As a result, by integrating precipitation information from MW-link, the estimated precipitation field can be further improved and the detailed validation is shown the next sub-section.

Furthermore, in Figure 10.3, it can also be found that the Copula parameters are mainly from the MW-link from Hohenpeissenberg to Weilheim, not the one from Hohenpeissenberg to Murnau. This may be due to the reason that the measurement quality of the MW-link from Hohenpeissenberg to Murnau is worse than the other one.

10.2.3 Validation

The results of the MW-link Copula parameter map included *Maximum Theta* approaches are checked using point wise cross-validation for all 31 stations, using the time series of gauge and pseudo observations in the respective grid cell. Different validation measures are used to obtain a quantitative appraisal for the performance of the simulations (see Table 10.1).

Generally the validation results are the same as shown in Table 7.2 of Chapter 7, the reason is because now only two MW-link derived Copula parameter maps can

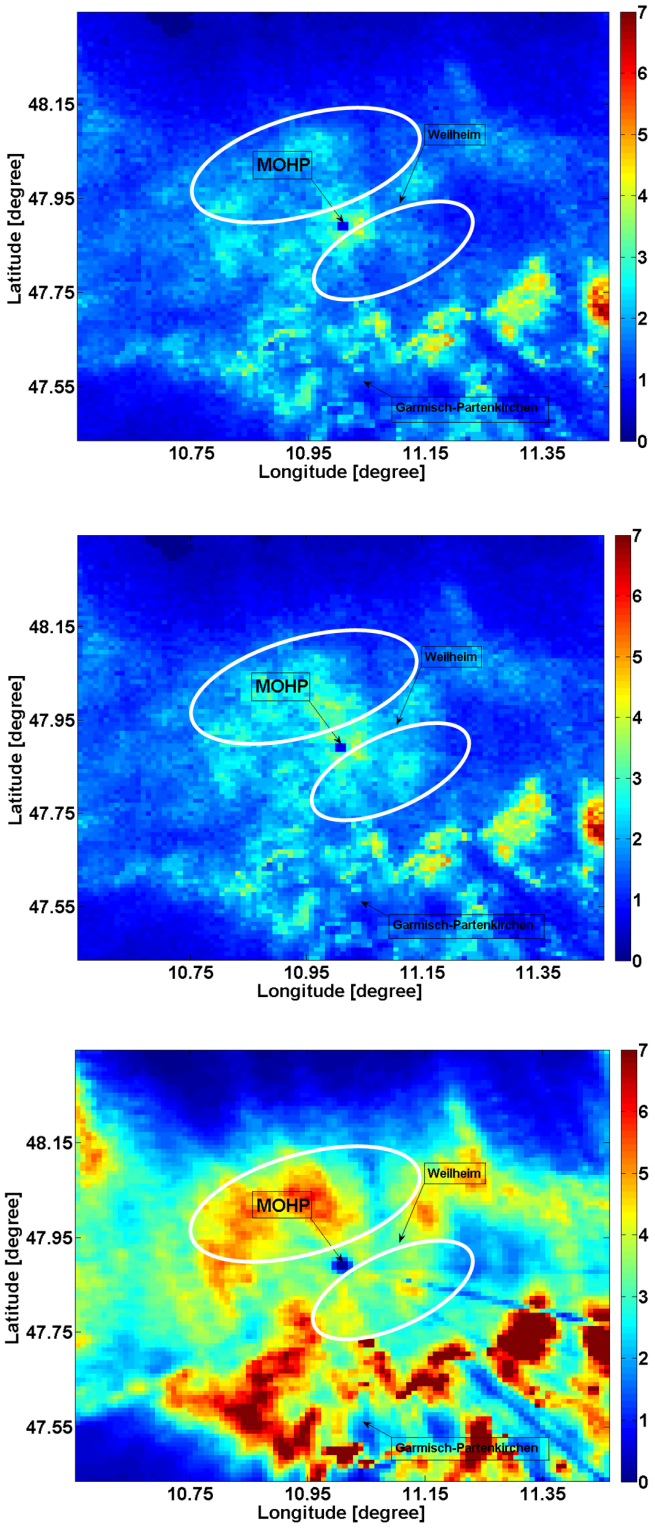

Figure 10.2: Pseudo observation field simulated by using *Maximum Theta* approach, w/o and w/ MW-link and original radar field (from top to bottom), 06:00, 04.08.2007, w/o ARMA-GARCH transformation. The unit of the color bar is in [mm/hour].

Table 10.1: Point wise cross-validation for MW-link included *Maximum Theta* approach in different stations, summer, 2008, w/o ARMA-GARCH transformation.

ID	Station Name	Kendall's τ [–]	r [–]	RMSE [mm/h]	MAE [mm/h]	NSE [–]
1	Bernbeuren-Prachtsried	0.64	0.76	1.70	0.96	0.23
2	Diessen	0.52	0.81	1.68	0.88	0.27
3	Deisenhofen	0.55	0.69	1.84	0.99	0.26
4	Ettal	0.55	0.70	1.61	0.98	0.35
5	Garmisch-Partenkirchen	0.54	0.64	1.74	0.92	0.31
6	Gilching	0.49	0.79	1.78	0.90	0.33
7	Griesen	0.51	0.70	1.46	0.84	0.39
8	Halblech	0.51	0.69	1.93	0.98	0.23
9	Hindelang	0.50	0.59	2.02	1.09	0.23
10	Hohenpeissenberg	–	–	–	–	–
11	Kaufbeuren	0.56	0.75	1.23	0.78	0.40
12	Kochel	0.57	0.74	1.58	0.91	0.30
13	Kohlgrub, Bad	0.62	0.80	1.75	0.97	0.32
14	Kraftisried	0.49	0.69	1.93	1.08	0.21
15	Kreuth	0.52	0.68	1.95	1.11	0.27
16	Krün	0.51	0.69	1.92	1.02	0.36
17	Lenggries	0.53	0.68	1.74	0.96	0.38
18	Maisach	0.45	0.74	1.60	0.91	0.33
19	Marktoberdorf	0.62	0.75	1.86	1.08	0.29
20	München	0.52	0.63	1.57	0.95	0.24
21	Oberammergau	0.58	0.76	1.59	0.89	0.35
22	Oberschleissheim	0.51	0.71	1.93	0.96	0.25
23	Oy	0.53	0.61	1.68	1.03	0.29
24	Schwangau	0.53	0.61	1.68	0.90	0.31
25	Seeg	0.57	0.67	2.01	1.06	0.15
26	Schäftlarn	0.59	0.78	1.71	0.95	0.18
27	Steingaden	0.69	0.89	1.90	0.93	0.41
28	Schwaben	0.51	0.55	2.00	1.05	0.12
29	Schlehdorf	0.59	0.74	2.03	1.05	0.25
30	Vilgertshofen	0.59	0.70	1.84	0.95	0.29
31	Wielenbach	0.55	0.72	**1.63**	**0.84**	**0.32**

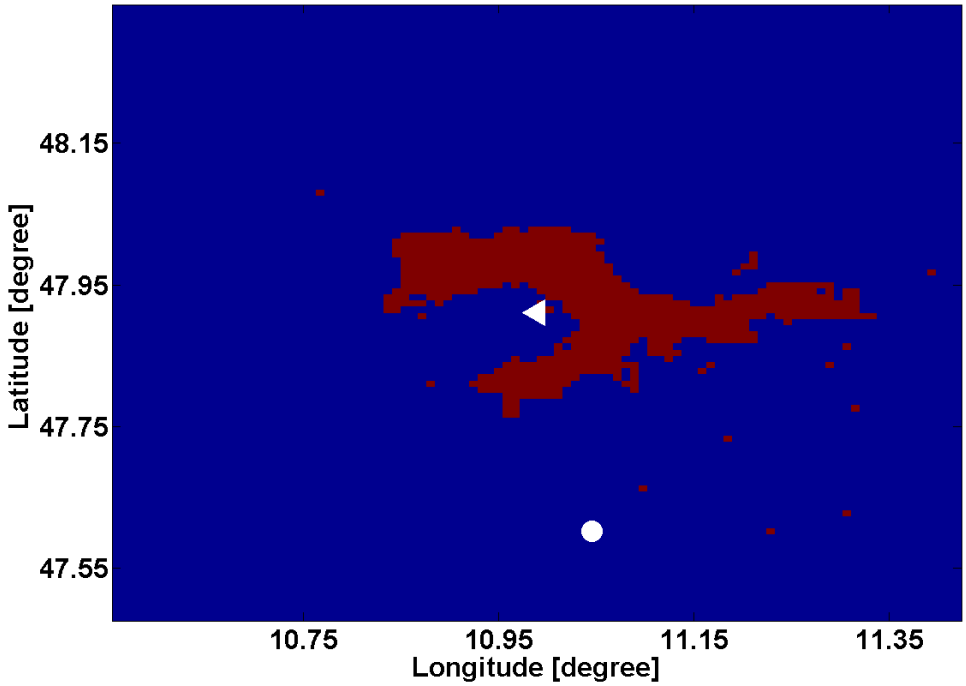

Figure 10.3: The area (grids highlighted with red) with the Copula parameters selected from the 2 MW-link derived Copula parameter maps at 06:00, 04.08.2007.

be used. This has been already mentioned in the last sub-section in Figure 10.3 that only the red areas can be improved by using the precipitation information from the 2 MW-links. However, small improvements can still be found in the two stations at Wielenbach, highlighted in bold, which are located not far away from the MW-link - Hohenpeissenberg to Weilheim. So, with more MW-links can be used in this region, it can be promising that the estimated rainfall fields can be improved.

10.3 MW-Link Based Adjustment

In this section, the approach is developed to use the precipitation information observed by the MW-link as the additional constrain to perform adjustment on the radar measurements (the first try or experiment due to the data availability problem).

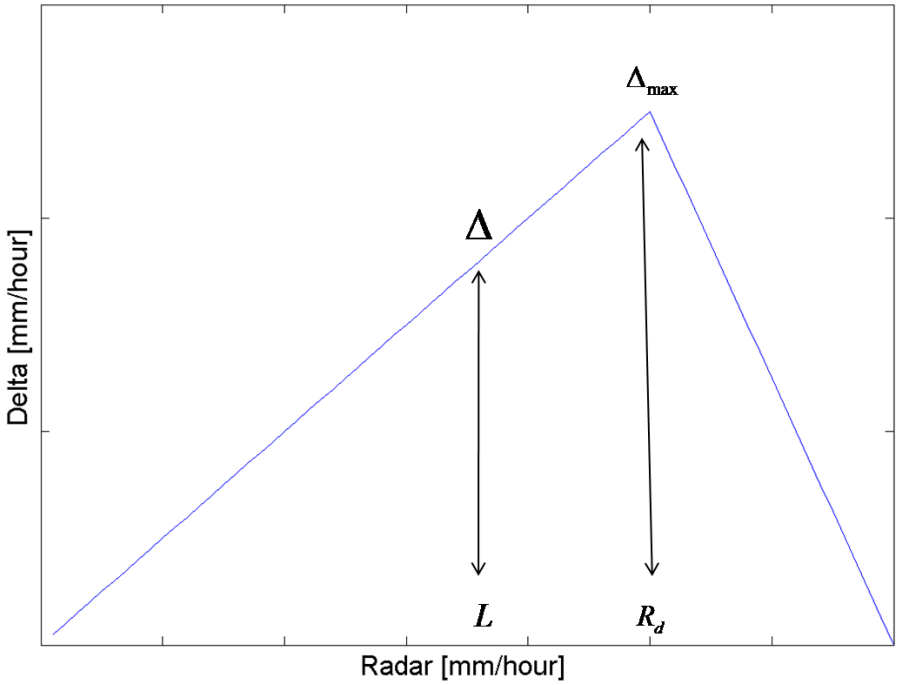

Figure 10.4: Theoretical representation of the distribution of δ_i based on the the corresponding measurements of R_i and the MW-link L, as well as the R_d.

10.3.1 Point Wise Adjustment

The difference Δ between the measurement (L) of one specific MW-link and the corresponding mean (R_{mean}) of the radar observations (e.g. $R_i, i = 1, ...m_1$) beneath this MW-link can be calculated as follows:

$$\Delta = L - R_{mean}. \tag{10.1}$$

This Δ is the mean or integrated difference between the MW-link and the corresponding radar measurements along this MW-link. In order to identify the individual $\delta_i, i = 1, ...m_1$ for each grid cell beneath this MW-link, the spatial distribution of this Δ has to be assumed beforehand. As suggested e.g. by AghaKouchak et al. 2010a and 2010b, the error or uncertainty for the radar observation can be assumed to be proportional to (or dependent on) the radar observation itself. Therefore, in this approach,

it is assumed that δ_i is dependent on the corresponding measurements from radar R_i and the MW-link L, as shown in the following formulae and Figure 10.4. The value of 25 mm/hour is chosen due to the data preprocessing steps as listed in Chapter 4.

$$\delta_i = \begin{cases} \Delta R_i / L & \text{if } R_i \leq R_d \\ \Delta(25 - R_i)/(25 - L) & \text{if } R_i > R_d \end{cases}$$

Here, the problem becomes how to determine the value for R_d, which lies on the principle that the mean of the sum of δ_i should be equal to Δ.

Afterwards, for each time step, the adjustment term δ_i can be added to the corresponding radar measurement R_i. In this way, the radar measurement is adjusted by the MW-link in the term of $R_i + \delta_i$. The basic steps to perform this point wise radar adjustment by using MW-link can be summarized as follows:

1. For each time step, the Δ is calculated between the MW-link observation L and the corresponding mean (R_{mean}) of the radar observations beneath this MW-link.

2. Following the assumption about the adjustment term, δ_i is calculated for each grid cell beneath this MW-link.

3. Finally, the MW-link adjusted radar measurement is calculated as $R_i + \delta_i$.

Figure 10.5 shows the Δ Vs resulting δ_i at the location of Hohenpeissernberg, marked as the blue circle, July to October, 2010. Meanwhile, as marked with the red circle, the differences between radar and gauge at Hohenpeissernberg are also presented. Their good match also implies that this assumption about the spatial distribution of δ_i is reasonable.

During the test period - July to October, 2010, by performing this MW-link (by using the MW-link from Hohenpeissenberg to Weilheim) based adjustment approach, the adjusted and non-adjusted radar observations are both compared to the rain gauge positive observations at Hohenpeissenberg as shown in Figure 10.6. It can be found that the adjusted radar measurements have a better match with the gauge observations than the original radar. Especially for the extreme large or small precipitation values, for instance at the time step 233 (236), large under-estimation (over-estimation) for original radar measurement is corrected by integrating the MW-link derived rainfall.

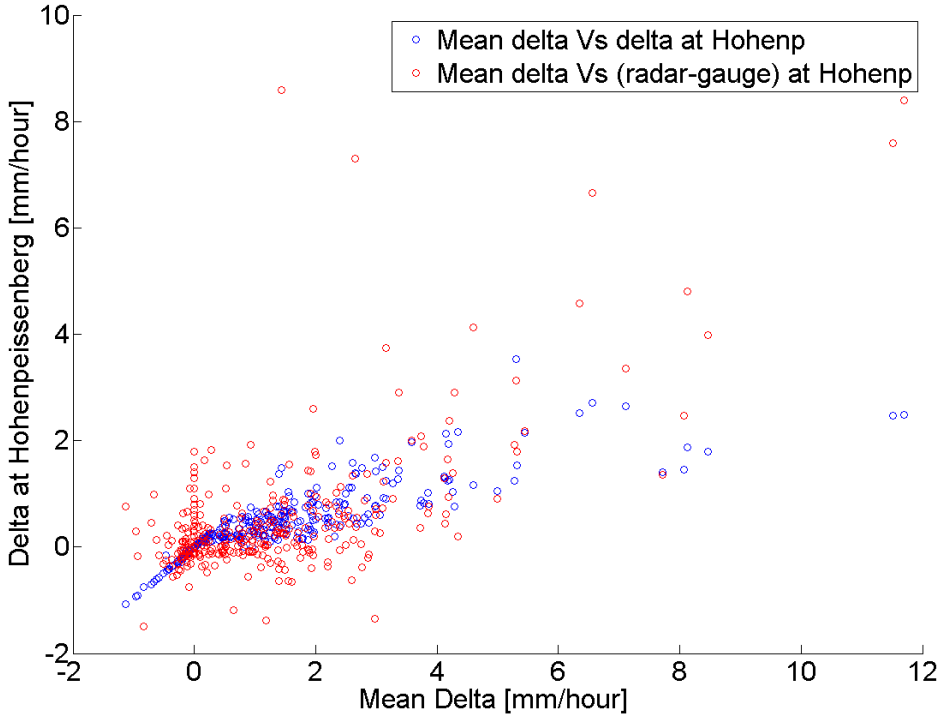

Figure 10.5: Δ Vs δ_i (blue circle) and radar gauge differences (red circle) at the location of Hohenpeissernberg, July to October, 2010.

Additionally, according to the quantity validation results by using Pearson's correlation and RMSE (both compared to the rain gauge observed positive series), it is further confirmed that the improvement by using this MW-link based adjustment approach is clear, 0.81 to 0.85 for Pearson's correlation and 1.41 to 1.05 mm/hour for RMSE, achieving the 5 percent and 25 percent better estimations (for the positive pairs only). Note that the radar measurements at neighbouring grid are used to represent for the observations at Hohenpeissenberg due to the radar data quality (same as in the previous chapters).

10.3.2 Spatial Extension

Since the adjustment term δ can be calculated from the given radar and MW-link measurements, the spatial extension for this MW-link based adjustment approach is straightforward. Once the radar measurement is known for one grid cell, the adjustment

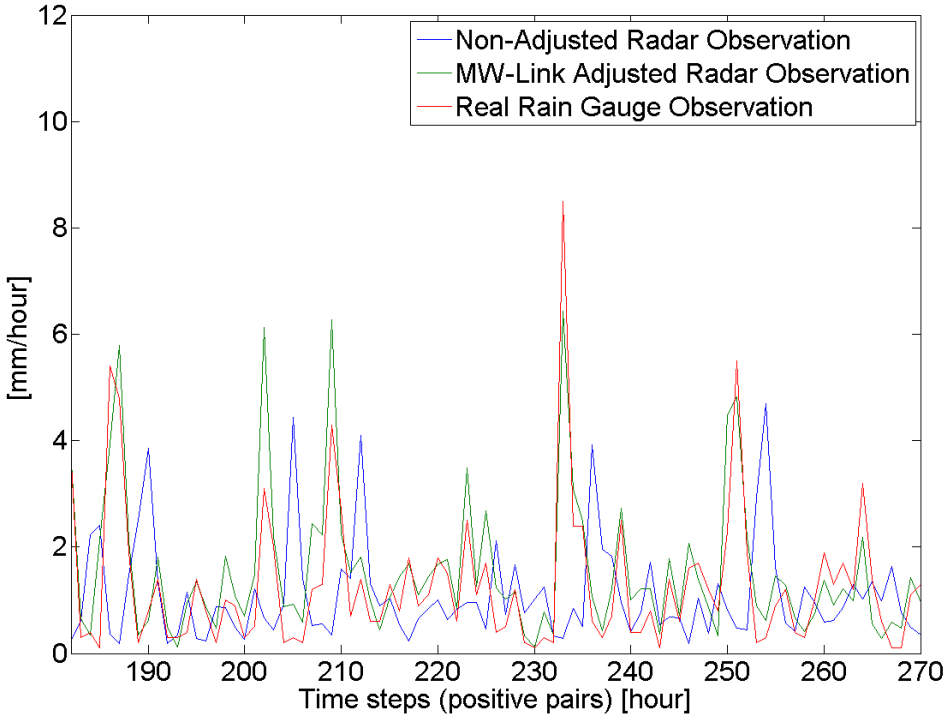

Figure 10.6: The positive time series for original and MW-link adjusted radar measurements, as well as the rain gauge observation at Hohenpeissenberg, July to October, 2010.

factor δ for this grid cell can be calculated immediately. Then, the adjusted observation can be generated by adding this adjustment factor δ to the original radar observation R.

Taking as one example, at Wielenbach, this MW-link based adjustment approach is applied and compared to the gauge observations at this grid cell as shown in Figure 10.7

It can be seen that, compared to the rain gauge observations at Wielenbach, the correlation for both original and MW-link adjusted radar measurements are equally good, but significant even larger over-estimations are found for MW-link adjusted ones. In fact, the original radar observations are already well matched with the gauge measurements at Wielenbach, with the Pearson's correlation coefficient at 0.90 and RMSE at 0.84 mm/hour (for positive pairs only). After performing the MW-link based adjust-

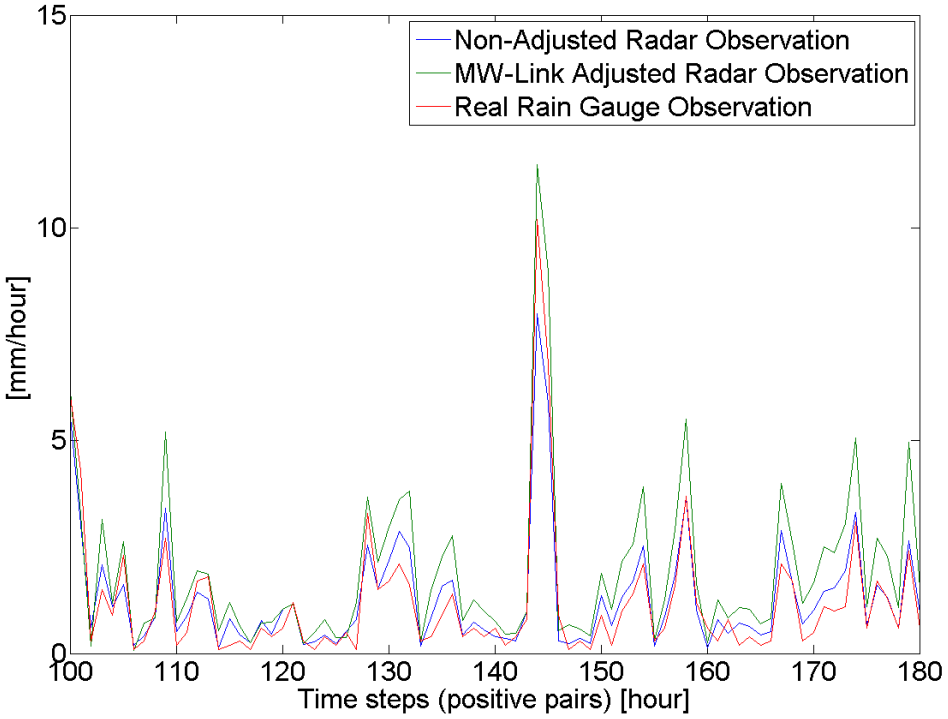

Figure 10.7: The positive time series for original and MW-link adjusted radar measurements, as well as the rain gauge observation at Wielenbach, July to October, 2010.

ment approach, a small improvement for Pearson's correlation is achieved from 0.90 to 0.92. However, for RMSE, it becomes even worse, reduced to 1.0 mm/hour, which may be due to the reasons as follows:

1. Wielenbach is not close to the MW-link from Hohenpeissenberg to Weilheim, about 5 km away from the Weilheim (north-eastern direction), so that the estimation results become even worse by including this MW-link.

2. The quality of the MW-link (from Hohenpeissenberg to Weilheim) derived rainfall is not as good as the radar measurements at Wielenbach (as shown for correlation and RMSE) so that the final results become worse when including the precipitation information from this MW-link.

Then it is important to define the effective range where the precipitation information derived from the MW-link can be used to perform the adjustments at spatial scale. The

MW-link derived Copula parameter map may provide one of the options. For instance, through setting up a threshold value of the Copula parameter, the grid cells with the Copula parameter larger than this threshold are defined where the MW-link based adjustment can be performed.

Or similar as the IDW, the MW-link derived Copula parameter map can also be used to replace the distance between two locations but using the ratio between their Copula parameter values. Taking the same example at Wielenbach, according to the generated Copula parameter map as shown in the top of Figure 10.1, the Copula parameter at Wielenbach CP_{Wie} and the mean value (CP_{mean}) of the Copula parameters along the grid cells beneath the MW-link (from Hoehenpeissenberg to Weilheim) can be identified. Then the ratio term CP_{Wie}/CP_{mean} can be calculated. So, the updated adjustment term $\delta_{Wie} + CP_{Wie}/CP_{mean}$ can be added to the original radar measurement.

In this case, by including the ratio CP_{Wie}/CP_{mean}, an improvement is achieved. By comparing to the gauge observations at Wielenbach, the Pearson's correlation is 0.91 and the RMSE is 0.90, 10 percent better than the results without introducing the ratio term.

This ratio term can also be used to integrate the precipitation information from other MW-links. For instance, there are l MW-links and then the adjusted term $\delta_i, i = 1, ..., l$ calculated from each MW-link can be integrated to composite the final adjustment term δ_{final} for one grid cell as listed below:

$$\delta_{final} = \sum_{i=1}^{l} \delta_i (CP/CP_{mean_i}) / \sum_{i=1}^{l} (CP/CP_{mean_i}). \tag{10.2}$$

So, the final adjusted radar measurement $R_{adjusted}$ for one grid cell is

$$R_{adjusted} = R_{non-adjusted} + \delta_{final}. \tag{10.3}$$

Furthermore, between this MW-link adjusted radar fields and rain gauge observations, the dependence structure can be captured by using the Copula-based approaches and then the pseudo observations can be generated based on the given MW-link adjusted radar measurements. Finally, the Copula-based approaches developed in this study can be directly used to generate the final precipitation fields assimilating data from radar, rain gauge and MW-link.

10.4 Summary and Discussion

In this chapter, at first, the precipitation information from MW-link attenuation is assimilated to estimate the rainfall field, based on the Copula parameter map based algorithm introduced in Chapter 7. Through MW-link derived Copula parameter maps which can be considered as the combination of several gauge derived Copula parameter maps, the dependence structures independent of gauge/radar are further integrated. Then the simulated field of pseudo observation can be produced by using the *Maximum Theta* approach combine data from gauge, radar and MW-link attenuation.

Through comparisons with the simulation results only with gauge/radar, some more detailed information can be identified with the additional rainfall measurements from MW-links, as well as the improvements confirmed in the validation part. Some problems still exists, i.e. the radar data availability so that only 2 MW-links can be used in this study. However, with the development and more widely use of this new remote sensing technology, more MW-links can be easily integrated in this Copula parameter map based data assimilation to estimate precipitation field, together with gauge and radar.

To including MW-link derived Copula parameter maps directly in the *Maximum Theta* approach is also a disadvantage with the indirect assumption that the MW-link is equivalent to the rain gauge. Therefore, it is still a challenge for how to integrate the Copula parameter maps. With the increasing number of MW-links, it is possible to generate precipitation fields only by using Copula parameter maps from MW-link and then to estimate the final rainfall field by further assimilating rain gauge observations.

In the second part, the MW-link based adjustment approach is further developed, which is somehow out of the scope of the Copula-based analysis. However, it is proved that with including this MW-link based adjustment, the resulted rainfall can be improved especially at the means of RMSE to correct the existed under or over-estimations. This adjustment method can also be performed to the Copula-based simulated pseudo observations to further correct the bias due to the rain gauge marginal distributions (i.e. to adjust the impacts from the relative humidity).

In this study, only the combination of the linear relationship is used to model the the spatial distribution of Δ. Of course, this adjustment approach can be further optimized by the better assumptions about the spatial distribution of Δ which requires more observations from radar, gauge and MW-link.

11

Summary, Conclusions and Perspectives

11.1 Summary and Conclusions

For estimating precipitation fields based on the observations, rain gauges are used as true ground measurements at point scale while the radar precipitation fields can additionally provide the spatial distribution. Nevertheless, precipitation fields from radar reflectivities incorporate significant errors due to measurement errors such as e.g. backscatter, shading, the used Z/R-relationship, and especially due to the physical principle that radar does not measure rainfall at the ground level (e.g. Vogl et al., 2012; Ehret et al., 2002). In this study, two Copula-based stochastic approaches are proposed to simulate the precipitation fields by assimilating rainfall observations from gauge and radar, through exploring their dependence structures. As a pre-requisite for the Copula-based analysis, the sensitivity from ARMA-GARCH transformation to the final precipitation fields is investigated. Additionally, the new rainfall observation devices - the MW-links - are also integrated to improve the estimated precipitation fields. As a summary, the following aspects are developed and discussed in this study:

1. As the foundation for the spatial simulation, the Copula-based point wise data integration approach is used to combine precipitation information from gauge and radar.

2. The statistical Copula parameter map based spatial scale data assimilation approaches, namely the *Multiple Theta* and the *Maximum Theta*, are developed to

produce the simulated fields of pseudo observation based on the observations from gauge and radar.

3. Furthermore, combining both physical and statistical points of view and including the impacts from range to the radar station and altitude, the interpolated Copula parameter field based data assimilation approach is developed to simulate precipitation fields of pseudo observation assimilating different data sources.

4. The simulation results w/o and w/ ARMA-GARCH transformation are compared at both the point and spatial scale to investigate its impact and sensitivity.

5. The different spatial distributions for rain gauge marginal distribution are further included in the Copula-based approached and their impacts on the simulation results are analysed and discussed.

6. The precipitation information from the MW-links is included to further improve to precipitation fields.

With the basic principles to trust the radar fields for the spatial rainfall pattern and to trust the rain gauges for the absolute values, those Copula-based rainfall field estimation approaches developed in this study are proven to have the capability to estimate precipitation fields combining the advantages from various observation sources. The approaches based on the Copula parameter map tend to incorporate anisotropies in the dependence structure. In the interpolated Copula parameter field based approach, the impacts from the range and local altitude are integrated and performed better than the Copula parameter map since the interpolated Copula θ field approach considering not only the statistical aspects but also the physical principles.

According to the GoF tests, the one parameter Gumbel Copula shows the best fit and is used to model the dependence structure between positive radar/gauge pairs. However, for the other Copula families with more than one parameters, it is difficult to perform the spatial simulations for both the Copula parameter map (*Maximum Theta*) and the interpolated Copula parameter field based approaches, the same for the temperature and altitude driven marginal distribution for rain gauges. However, the *Multiple Theta* approach is not restricted and can be applied also for more than one theoretical Copula family and Copulas with multidimensional parameter spaces (Vogl et al., 2012).

For different seasons different types of rainfall regimes are predominant (Vogl et al., 2012). Therefore, to generate Copula parameter maps and interpolated Copula parameter fields for different seasons is expected to reveal preferential precipitation types. Furthermore, the temperature dependent radar/gauge dependence structure (see 9.6) has the promising ability to to realise the dynamic Copulas (different Copula parameter for different temperatures).

Another possible solution is the temperature and altitude driven marginal distribution for rain gauges. Generally speaking, besides the estimation of the Copulas, the marginal distribution for rain gauges is another key factor for the Copula-based simulation. Through involving the variables such as temperature and relative humidity, more detailed physical backgrounds of the precipitation process are included in the Copula-based rainfall fields estimation.

Theoretically, the best way is to use multivariate Copula to model the dependence structure among precipitation from gauge, radar, temperature, relative humidity and etc., in one Copula function. However, the existing multivariate Copula families, such as Gaussian and student-T are not performing well enough to present those sophisticated multi-dimensional dependence structures, particularly for hydro-meteorological variables. The development of more efficient multivariate Copula family is an urgent task. The v-transformed Copula can be one option (Bárdossy and Li, 2008).

Considering the *i.i.d.* assumption, as a constraint for Copulas, it is necessary to perform the ARMA-GARCH transformation. In this study, it has been proved that, w/ ARMA-GARCH transformation, the dependence structure between gauge and the corresponding radar is reduced as demonstrated by the smaller estimated Copula parameters so that the simulation results become even worse. It is worth to note that, in the financial field (where the Copulas and ARMA-GARCH transformation come from), all the data are not affected by the measurement errors. In contrast, as known, large errors or uncertainties are existing in the radar measurements and also for rain gauges. These errors or uncertainties between radar and rain gauge should not be correlated due to their different measurement principles. On the other hand, in their temporal structures, there should be no autocorrelation structures. Therefore, after ARMA-GARCH transformation, those errors or uncertainties may still be remained in the residuals so that the Copula parameters estimated from the residuals are smaller than those from the original data sets. In order to prove this assumption, for instance, the dummy data

set can be developed with the objectively built errors and to be tested if those errors are removed after the ARMA-GARCH transformation.

It is necessary to identify the error or uncertainty beforehand. In order to do this, the simple and efficient way is to have several rain gauges in one radar grid cell and then use the mean value of those rain gauges as the reference observation. However, in this study, no such grid can be found. In another work done by AghaKouchak et al. (2010a and 2010b), a Copula-based method is presented to simulate radar error fields using a network of gauge stations as reference by using different multivariate Copula distributions but with the shortcoming of extensive computation cost (Vogl et al., 2012). Instead, the MW-links are proven to be capable to adjust the radar fields at first and then the Copula-based analysis can be performed to generate the precipitation fields assimilating data from radar, gauge and MW-link.

11.2 Perspective of Future Work

Based on the analysis carried out in this study, the results discussed in the previous chapters, and the scope of this research, the following recommendations for future research are suggested:

1. Overcoming the problem that only the mean value of 100 repeated simulation is used as the final estimation, but to identify the exact simulation value among the conditional PDF behind.

2. Exploring more functional spatial distribution for rain gauge marginal distributions, for instance to include the impact from relative humidity within the temperature and altitude dependent approach.

3. Improving the back-transformation, especially for the events with high rainfall intensities in the rank space but resulting with the low intensities in the data space after the back-transformation by using the estimated rain gauge marginal distributions.

4. Applying better methods to perform the rainfall types classification, e.g. to use the vertical radar volume scanning product.

5. Comparing precipitation fields estimated by using the Copula-based approaches to other rainfall fields, for instance the REGNIE field, or rainfall fields derived from traditional interpolation or assimilation methods such as co-kriging and etc..

6. Involving more available MW-links with the possibility to analyse the error structure or distribution for both radar and rain gauges, not only temporal but also on the spatial scale.

Bibliography

[1] Abdi, H., and Williams, L.J.: Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2: 433459, 2010.

[2] AghaKouchak, A., Bárdossy, A., and Habib, H.: Copula-based uncertainty modeling: Application to multi-sensor precipitation estimates. Hydrological Processes 24, doi:10.1002/hyp.7632, 2010a.

[3] AghaKouchak, A., Bárdossy, A., and Habib, H.: Conditional simulation of remotely sensed rainfall data using a non-gaussian v-transformed copula, Adv. Water Resour., 33, 624634, doi:10.1016/j.advwatres.2010.02.010, 2010b.

[4] AghaKouchak, A., Habib, E., and Bárdossy, A.: Modeling radar rainfall estimation uncertainties: random error model, J. Hydrol. Eng., 15, 265275, 2010c.

[5] Ajayi, G.O., and Kozu, T.: Rain drop size distribution during convective rainfall in Japan and Nigeria. Proc. of Abstracts, GA URSI, FP37, 1999.

[6] Akaike, H.: A new look at the statistical model identification. IEEE Transactions on Automatic Control 19 (6): 716723. doi:10.1109/TAC.1974.1100705. MR0423716, 1974.

[7] Bárdossy, A.: Copula-based geostatistical models for groundwater quality parameters. Water Resources Research, 42(11), W11416.1W11416.12, doi:10.1029/2005WR004754, 2006.

[8] Bárdossy, A., and Li, J.: Geostatistical interpolation using copulas. Water Resources Research, 44(7), W07412.1W07412.15, doi:10.1029/2007WR006115, 2008.

[9] Bárdossy, A., and Pegram, G.G.S.: Copula based multisite model for daily precipitation simulation. Hydrology and Earth System Sciences, 13, 2299-2314, 2009.

[10] Bartels, H., Weigl, E., Reich, T., Lang, P., Wagner, A., Kohler, O., and Gerlach, N.: Projekt RADOLAN, Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer), Abschlussbericht (in German), Deutscher Wetterdienst, 2004.

[11] Bates, B.C., Kundzewicz, Z.W., Wu, S., and Palutikof, J.P., Eds.: Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva, 210 pp, 2008.

[12] Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327, doi:10.1016/0304-4076(86)90063-1, 1986.

[13] Bouilloud, L.: Radar rainfall estimation in the context of post-event analysis of flash-flood events. J. Hydrol, doi:10.1016/j.jhydrol.2010.02.035, 2010.

[14] Bourgouin P.: A Method to Determine Precipitation Types. Bull. Amer. Meteor. Soc., 15, 583592, 2000.

[15] Box, G.E., Jenkins, G.M., and Reinsel, G.C.: Time series analysis: forecasting and control, 3rd edition, Prentice Hall Englewood Cliffs, New Jersey, 1994.

[16] Brandes, E.A.: Optimizing Rainfall Estimates with the Aid of Radar. J. Appl. Meteor., 14, 1339-1345, 1975.

[17] Burrough, P.A., and McDonnell, R.A.: Principles of Geographical Information Systems. Oxford University Press, Oxford, 1998.

[18] Chwala, C., Siart, U., Hipp, S., Eibert, T., and Kunstmann, H.: Dual-band polarimetric coherent atmospheric transmission experiment for precipitation observation. Microwave Conference (EuMC), 2010 European, pp. 13331336, 2010.

[19] Chwala, C., Gmeiner, A., Qiu, W., Hipp, S., Nienaber, D., Siart, U., Eibert, T., Pohl, M., Seltmann, J., Fritz, J., and Kunstmann, H.: Precipitation observation using microwave backhaul links in the alpine and pre-alpine region of Southern Germany, Hydrol. Earth Syst. Sci. Discuss., 9, 741-776, doi:10.5194/hessd-9-741-2012, 2012.

[20] Cole, S.J., and Moore, J.: Hydrological modelling using rain gauge and radar based estimators of areal rainfall. J. of Hydrol., 358, 159-181, 2008.

[21] Collier, C.: Accuracy of Rainfall Estimates by Radar, Part I: Calibration by Telemetering Rain-gauges. Journal of Hydrology, 83, pp. 207 223, 1986.

[22] Collins, F.B., and Bolstad, P.V.: A comparison of spatial interpolation techniques in temperature estimation, Proceedings, Third International Conference/Workshop on Integrating GIS and Environmental Modelling, Santa Fe, NM. Santa Barbara, CA: National Center for Geographic Information and Analysis, Santa Barbara, 1996.

[23] Cummings, R.J., Upton, G.J G., Holt, A R., and Kitchen, M.: Using microwave links to adjust the radar rainfall field. Advances in Water Resources. Volume: 32, Issue: 7, Publisher: Elsevier Ltd, Pages: 1003-1010, 2009.

[24] Dupuis, D.: Using copulas in hydrology: Benets, cautions, and issues. Journal of Hydrologic Engineering, 12(4), 381393, 2007.

[25] DWD: RADOLAN − Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer). Zweiter Zwischenbericht, Deutscher Wetterdienst, Geschftsfeld Hydrometerologie, 12/98, 1998.

[26] DWD: RADOLAN − Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer). Vierter Zwischenbericht, Deutscher Wetterdienst, Geschftsfeld Hydrometerologie, 01/00, 2000.

[27] DWD: RADOLAN − Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer). Fnfter Zwischenbericht, Deutscher Wetterdienst, Geschftsfeld Hydrometerologie, 04/01, 2001.

[28] Ehret, U.: Rainfall and Flood Nowcasting in Small Catchments using Weather Radar. PhD dissertation, ISBN 3-933761-24-7, 2002.

[29] Engle, R.: Autorregressive Conditional Heteroskedasticity with Estimates of United Kingdom Inflation. Econometrica, 50, 987-1008, 1982.

[30] Engle, R.: Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica. Vol. 96, pp. 893920, 1988.

[31] Erxleben, J., Elder, K., and Davis, R.: Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains. Hydrological Processes, 16: 3627-3649, 2002.

[32] Favre, A.C., El Adlouni, S., Perreault, L., Thiémonge, N., and Bobee, B.: Multivariate hydrological frequency analysis using copulas. Water Resources Research 40, 112, 2004.

[33] Fiser, O.: A simple generator of forward scattering functions on spherical dielectrics. Radio Engineering, 2(1), 21-22, 1993.

[34] Fiser, O.: The Role of DSD and Radio Wave Scattering in Rain Attenuation, Geoscience and Remote Sensing New Achievements, Pasquale Imperatore and Daniele Riccio (Ed.), ISBN: 978-953-7619-97-8, In Tech, 2010.

[35] Frees, E.W., and Valdez, E.A.: Understanding relationships using copulas. North American Actuarial Journal, 2(1), 1-25, 1998.

[36] Genest, C., and Rivest L.-P.: Statistical inference procedures for bivariate Archimedean copulas. J. Amer. Statist. Assoc. 88, 10341043, 1993.

[37] Genest, C., and Favre, A.-C.: Everything you always wanted to know about copula modeling but were afraid to ask. Journal of Hydrologic Engineering 12 (4), 347-368, 2007.

[38] Genest, C., and Remillard B.: Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. Annales de l'Institut Henri Poincare: Probabilites et Statistiques, 44, 1096-1127, 2008.

[39] Genest, C., Remillard, B., and Beaudoin, D.: Goodness-of-fit tests for copulas: A review and a power study. Insurance: Mathematics and Economics, 44, 199-214, 2009.

[40] Giuli, D., Toccafondi, A., Bif Gentili, GA., and Freni, A.: Tomographic reconstruction of rainfall elds through microwave attenuation measurements. J. Appl. Meteorol., 1991;30(9):132340, 1991.

[41] Giuli, D., Facheris, L., and Tanelli, S.: Microwave tomographic inversion technique based on stochastic approach for rainfall elds monitoring. IEEE Trans Geosci Remote Sens 1999;37(5):253655, 1999.

[42] Giuli, D., Facheris, L., and Tanelli, S.: Microwave tomographic inversion technique based on a stochastic approach for rainfall fields monitoring, IEEE Trans. Geo. Rem. Sens., vol. 37(5), December, pp 2536-2555, 2007.

[43] Gomez-Hernandez, J., and Wen, X.: To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. Adv. Water Resour., 21, 47-61, 1998.

[44] Goovaerts, P.: Performance comparison of geostatistical algorithms for incorporating elevation into the mapping of precipitation. Prcoceedings IV International Conference on GeoComputation, Fredericksburg, VA, USA, 25.-28. July, 1999.

[45] Goovaerts, P.: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J. Hydrol., 228, 113-129, 2010.

[46] Grégoire, V., Genest, C., and Gendron, M.: Using copulas to model price dependence in energy markets. Energy Risk, 5 (5), 58-64, 2008.

[47] Habib, E., Aduvala, A., and Meselhe, E.: Analysis of radar rainfall error characteristics and implications for streamflow simulations uncertainty. J. Hydrologic Sciences, 2008.

[48] Haberlandt, U.: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. J. Hydrol., 332, 144-157, 2007.

[49] Holt, A., Kuznetsov, G., and Rahimi, A.: Comparison of the use of dual-frequency and single-frequency attenuation for the measurement of path-averaged rainfall along a microwave link, Microwaves, Antennas and Propagation, IEE Proceedings -, 150, 31520, doi:10.1049/ip-map:20030616, 2003.

[50] Hutchinson, M.F.: Interpolation of rainfall data with thin plate smoothing splines: I Two dimensional smoothing of data with short range correlation. Journal of Geographic Information and Decision Analysis 2 (2), 139151, 1998a.

[51] Hutchinson, M.F.: Interpolation of rainfall data with thin plate smoothing splines: II Analysis of topographic dependence. Journal of Geographic Information and Decision Analysis 2 (2),152167, 1998b.

[52] van de Hulst, H.C.: Light Scattering by Small Particles. New York: Wiley, 1957.

[53] IPCC: Impacts, Adaptation, and Vulnerability. Exit EPA Disclaimer Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Parry, Martin L., Canziani, Osvaldo F., Palutikof, Jean P., van der Linden, Paul J., and Hanson, Clair E. (eds.)]. Cambridge University Press, Cambridge, United Kingdom, 2007.

[54] Ishimaru, A., and Lin, J.C.: Multiple scattering effects on wave propagation through rain, in NATO/AGRARD conf.Proc., No. 107, pp. 1-13, North Atlantic Treaty Organisation, Brussels, Belgium, 1973.

[55] Isaaks, E.H., and Srivastava, R.M.: Applied Geostatistics. Oxford University Press, New York, 1989.

[56] ITU: RECOMMENDATION ITU-R P.838-2 from the International Telecommunication Union: Specific attenuation model for rain for use in prediction methods, 2003.

[57] Joe, H.: Multivariate Models and Dependence Concepts. Chapman and Hall, New York, 1997.

[58] Joss, J., and Lee, R.: The application of radar-gauge comparisons to operational precipitation prole corrections. J. Appl. Meteorol., 34, 2612263x, 1995.

[59] Krajewski, W.F.: Cokriging radar-rainfall and rain-gauge data. Journal of Geophysical Research: Atmospheres 92 (D8), 9571-9580, 1987.

[60] Krajewski, W.F., and Smith, J.: Radar hydrology: Rainfall estimation. J. Hydrol., 25, 13871394, 2002.

[64] Kun, Z., Wenzhong, G., and Guoqing, L.: The improved Kalman filter combined with Variational Analysis A new way to calibrate radar rainfall. Proceedings of 5th International Symposium on Hydrological Applications of Weather Radar, Heian-Kaian, Kyoto, Japan, November 19-22, 2001.

[62] Kuchment, L.S.: The Hydrological Cycle and Human Impact on It. Water Resources Management, Encyclopedia of Life Support Systems, 2004.

[63] Kuhn, G., Khan, S., and Ganguly, A.: Geospatial-temporal dependence among weekly precipitation extremes with applications to observations and climate model simulations in south america. Journal of Advances in Water Resources 30, 2401-2423, 2007.

[64] Kyriakidis, Phaedon C., Kim, Jinwon and Miller, Norman L.: Geostatistical Mapping of Precipitation from Rain Gauge Data Using Atmospheric and Terrain Characteristics. J. Appl. Meteor., 40, 18551877, 2001.

[65] Laux, P., Vogl, S., Qiu, W., Knoche, H.R., and Kunstmann, H.: Copula-based statistical refinement of precipitation in RCM simulations over complex terrain. Hydrol. Earth Syst. Sci., 15, 1-19, 2011.

[66] Leijnse, H., Uijlenhoet, R., and Stricker, J.N.M.: Hydrometeorological application of a microwave link: 2. Precipitation, Water Resources Research, 43, 9 PP., doi:200710. 1029/2006WR004989, 2007a.

[67] Leijnse, H., Uijlenhoet, R., and Stricker, J.N.M.: Rainfall measurement using radio links from cellular communication networks, Water Resources Research, 43, 6 PP., doi:200710. 1029/2006WR005631, 2007b.

[68] Leijnse, H., Uijlenhoet, R., and Stricker, J.N.M.: Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation, Advances in Water Resources, 31, 14811493, doi:10.1016/j.advwatres.2008.03.004, 2008.

[69] Mandapaka, P.V., Krajewski, W.J., Ciach, G.J., Villarini, G., and Smith, J.A.: Estimation of radar-rainfall error spatial correlation. Advances in Water Resources, 32, 1020-1030, 2009.

[70] Matsoukas, C., and Islam, S.: Fusion of radar and rain gage measurements for an accurate estimation of rainfall. Journal of Geophysical Research, Vol. 104, No. D24, 1999.

[71] Marx, A.: Einsatz gekoppelter Modelle und Wetterradar zur Abschtzung von Niederschlagsintensitten und zur Abflussvorhersage. PhD dissertation, ISBN 3-933761-64-6, 2007.

[72] Messer, H., Zinevich, A., and Alpert, P.: Environmental monitoring by wireless communication networks. Science, 312, (5774), 713, 2006.

[73] Michele, C.De, and Salvadori, J.: A generalized pareto intensity-duration model of storm rainfall exploiting 2-copulas. J. Geophy. Res. Atmos. 108 (D2), 111, doi:10.1029/2002JD002534, 2003.

[74] Mie, G.: Beitrge zur Optik trber Medien, speziell kolloidaler Metallsungen, Annalen der Physik, Vierte Folge, 25(3), pp. 377-445, 1908.

[75] Minda, H., and Nakamura, K.: High Temporal Resolution Path-Average Rain Gauge with 50-GHz Band Microwave, Journal of Atmospheric and Oceanic Technology, 22, 165179, 2005.

[76] Moore, R., May, B., Jones, D., and Black, K.: Local Calibration of Weather Radar Over London. Advances in Radar Hydrology, European Commission, 1994b.

[77] Nelsen, R.B.: An Introduction to Copulas. Springer-Verlag, New York, 1999.

[78] Olsen, Roderic L., Rogers, David V., and Hodge, Daniel B.: The aR^b Relation in the Calculation of Rain Attenuation, IEEE Transactions on Antennas and Preopagation, Vol. AP-26, No. 2, pp 318-329, March 1978.

[79] Orr, S., Cartwright, A., and Tickner, D.: WWF Water Security Series 4 - Understanding water risks - A primer on the consequences of water scarcity for government and business, 2009.

[80] Pahl-Wost, C.: Transitions towards adaptive management of water facing climate and global change. Water Resource Manage, 21:4962 doi 10.1007/s11269-006-9040-4, 2007.

[81] Penndorf, R.B.: Scattering and extinction coefficients for small absorbing and non-absorbing aerosols. J.Opt. Sco.Amer., vol.52, pp.896-904, 1996.

[82] Rahimi, A.R., Upton, G.J.G., and Holt, A.R.: Dual-frequency links a complement to gauges and radar for the measurement of rain, Journal of Hydrology, 288, 312, doi:10.1016/j.jhydrol. 2003.11.008, 2004.

[83] Renard, B., and Lang, M.: Use of a gaussian copula for multivariate extreme value analysis: Some case studies in hydrology. Advances in Water Resources 30, 897-912, 2007.

[84] Rincon, R.F., and Lang, R.H.: Microwave link dual-wavelength measurements of path-average attenuation for the estimation of drop size distributions and rainfall, IEEE Transactions on Geoscience and Remote Sensing, 40, 760770, 2002.

[85] Roe, Gerard H.: Orographic Precipitation. Annu. Rev. Earth. Planet. Sci. 2005.33:645-671. doi: 10.1146/annurev.earth.33.092203.122541, 2005.

[86] Rogers, D.V., and Olsen, R.L.: Calculation of radiowave attenuation due to rain at frequencies up to 1000 GHz, Communications Res,Center, Dep. of Communications, Ottowa, ON, Canada, CRC Rep 1299, 1976.

[87] Salvadori, G., De Michele, C., Kottegoda, N.T., and Rosso, R.: Extremes in Nature: an approach using copulas. Springer-Verlag, New York, 2007.

[88] Schölzel, C., and Friederichs, P.: Multivariate non-normally distributed random variables in climate research introduction to the copula approach. Nonlinear Processes in Geophysics, 15(5), 761-772, 2008.

[89] Seo, D., Krajewski, W., and Bowles, D.: Stochastic interpolation of rainfall data from rain-gauges and radar using kogriging. 1. Design of experiments. Water Resources Research, 26(3), 469 477, 1990a.

[90] Seo, D., Krajewski, W., and Bowles, D.: Stochastic interpolation of rainfall data from rain-gauges and radar using kogriging. 2. Results. Water Resources Research, 26(5), 915 924, 1990b.

[91] Seo, D.: Rainfall Estimation in the Nexrad Era-Operational Experience, Issues, and Ongoing Efforts in the U.S. National Weather Service. Proceedings Second International River Basin modelling, management and flood mitigation (RIBAMOD), Montselice, Italy, published 1999.

[92] Serinaldi, F.: Analysis of inter-gauge dependence by kendall's τ, upper tail dependence coefficient, and 2-copulas with application to rainfall fields. Stochastic Environmental Research and Risk Assessment, 22, 671688, 2008.

[93] Singh, V.P.: Effects of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. Hydrological Processes, 11 (12), 1649-1669, 1997.

[94] Sklar, K.: Fonctions de repartition à n dimensions et leura marges. Publications de´ l Institut de Statistique de LUniversité de Paris 8, 229-231, 1959.

[95] Smith, J., Seo, D., Baeck, M., and Hudlow, M.: An intercomparison study of NEXRAD precipitation estimates. Water Resources Research, 32(7),pp.2035-2045, 1996.

[96] Steiner, M., Houze Jr., R.A., and Yuter, S.E.: Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data, Journal of Applied Meteorology, 34, 19782007, 1995.

[97] Villarini, G., Serinaldi, F., and Krajewski, W.F.: Modeling radar-rainfall estimation uncertainties using parametric and non-parametric approaches. Advances in Water Resources, 31(12), 1674-1686, 2008.

[98] Vogl, S., Laux, P., Qiu, W., Mao, G., and Kunstmann, H.: Copula-based assimilation of radar and gauge information to derive bias corrected precipitation fields, Hydrol. Earth Syst. Sci. Discuss., 9, 937-982, doi:10.5194/hessd-9-937-2012, 2012.

[99] Wagner, A., Clemens, M., and Seltmann, J.: Vertical profile of drop size spectra, Proceedings of ERAD (2004): 402406c Copernicus GmbH, 2004.

[100] Wilson, J., and Brandes, E.: Radar measurement of rainfall A summary. Bulletin of the American Meteorological Society, Vol. 60, No. 9, pp. 1048 1055, 1979.

[101] Zhang, L.S.R., and Singh, V.: Bivariate rainfall frequency distributions using archimedean copulas. Journal of Hydrology, 32(1-2), 93-109, 2008.

[102] Zinevich, A., Alpert, P., and Messer, H.: Estimation of rainfall fields using commercial microwave communication networks of variable density. Advances in Water Resources, Vol: 31,Issue: 11, 1470-1480, 2008.

Declaration

I herewith declare that I have produced this paper without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This paper has not previously been presented in identical or similar form to any other German or foreign examination board.

The thesis work was conducted from December, 2008 to July, 2012 under the supervision of Prof. Dr. Harald Kunstmann both at University of Augsburg, Institute for Geography and KIT, IMK-IFU, Germany.

Wei Qiu, Garmisch-Partenkirchen, 30.07.2012

Curriculum Vitae

PERSONAL DATA

Name: Qiu Wei Sex: Male

Birthday: Jan 21st, 1983 Birthplace: Shengzhou, Zhejiang, China

E-mail: heinrich.qiu@gmail.com

EDUCATION BACKGROUND

Sept. 2001–Jun. 2006 **Bachelor of Science**

Electronic Science and Technology, University of Science and Technology of China, China.

Oct. 2006–Oct. 2008 **Master of Science**

Earth Oriented Space Science and Technology, Technische Universität München, Germany.

Dec. 2008–Aug. 2012 **Promotion**

Universität Augsburg & Karlsruhe Institute of Technology, IMK-IFU, Germany.

Dissertation: Copula-Based Precipitation Fields Estimation Combining Data from Radar, Gauge and Microwave Attenuation.

Supervisor: Prof. Dr. Harald Kunstmann